DOCUMENT RESUME

ED 415 641 EC 306 124

AUTHOR	Liu, Kristin; Thurlow, Martha; Erickson, Ronald; Spicuzza, Richard; Heinze, Kathryn
TITLE	A Review of the Literature on Students with Limited English Proficiency and Assessment. State Assessment Series: Minnesota, Report 11.
INSTITUTION	Minnesota State Dept. of Children, Families, and Learning, St. Paul.; National Center on Educational Outcomes, Minneapolis, MN.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	1997-08-00
NOTE	66p.
CONTRACT	R279A50011
AVAILABLE FROM	National Center on Educational Outcomes, University of Minnesota, 350 Elliott Hall, 75 East River Road, Minneapolis, MN 55455; phone: 612-626-1530; fax: 612-624-0879; World Wide Web: http://www.coled.umn.edu/NCEO (document may be copied without charge, additional print copies, $15).
PUB TYPE	Information Analyses (070)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Academic Standards; Accountability; *Achievement Tests; *Educational Assessment; Elementary Secondary Education; Evaluation Methods; Language Minorities; *Limited English Speaking; *Minimum Competency Testing; Outcomes of Education; Performance Factors; State Programs; Student Characteristics; *Student Evaluation; *Student Participation; Testing Problems
IDENTIFIERS	*Testing Accommodations (Disabilities)

ABSTRACT

In order to encourage school districts to include students with Limited English Proficiency (LEP) in educational assessments, this report discusses issues involved in assessing students with LEP. Topics include: (1) the role of testing in education and the different types of tests mandated by educational reform legislation; (2) characteristics of students with LEP and relevant demographic data; (3) factors within the student, the student's environment, and the testing environment that can influence the test results; (4) modifications and adaptations that make assessments more appropriate for students with LEP; (5) research findings related to current assessment practices in states and school districts across the country; and (6) research findings from large-scale assessment at the national level. Findings from the literature review are highlighted, including the trend toward large-scale assessments, the important role of testing, the growing population of students with LEP, the lack of a standardized definition for LEP, the need for clear guidelines for testing of students with LEP, the need to consider certain factors when language minority students and students with LEP are tested, and the need for states to pool their knowledge about students with LEP and their assessment resources. (Contains 70 references.) (CR)

ED 415 641

A Review of the Literature on Students with Limited English Proficiency and Assessment

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EC 306124

BEST COPY AVAILABLE

STATE ASSESSMENT SERIES
Minnesota Report 11

A Review of the Literature on Students with Limited English Proficiency and Assessment

Minnesota Assessment Project

Project Staff:
Constance Anderson • Cathy Wagner • Michael Tillman
Minnesota Department of Children, Families and Learning

Project Advisors:
Karon Hergenhahn • Boulieng Phommasouvanh • Carol Quest •
Marilyn Sorensen • Barbara Jo Stahl • Elizabeth Watkins • Nancy Larson
Minnesota Department of Children, Families and Learning

Prepared By:
Kristin Liu • Martha Thurlow • Ronald Erickson • Richard Spicuzza
University of Minnesota

Kathryn Heinze
Hamline University

August 1997

The Minnesota Assessment Project is a four-year, federally funded effort awarded to the Minnesota Department of Children, Families and Learning from the U.S. Department of Education, Office of Educational Research and Improvement. The project's goal is to promote and evaluate the participation of students with limited English proficiency and students with disabilities in Minnesota's Graduation Standards. Specifically, the project will examine ways in which students with limited English and students with disabilities can participate in the Basic Standards Exams of reading, mathematics and written composition and in the performance-based assessments of the high standards in the Profile of Learning.

This project is supported, in part, by a grant to the Minnesota Department of Children, Families and Learning from the U.S. Department of Education, Office of Educational Research and Improvement (Grant #R279A50011). Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

Permission is granted to copy this document without charge. Additional print copies may be purchased for $15.00 from:

NATIONAL CENTER ON EDUCATIONAL OUTCOMES
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612.626.1530 • Fax 612.624.0879
http://www.coled.umn.edu/NCEO

This document is available in alternative formats upon request.

High Standards and LEP Students

As a part of the national movement for all students in public school to achieve high standards, Minnesota has implemented the Minnesota Basic Standards Tests, which students must pass in order to receive a high school diploma. These high stakes tests measure minimum competencies in reading, math, and writing. Currently, there is controversy over the appropriateness of the tests for certain groups of students, in particular for limited English proficient (LEP) students. Some educators believe that since the tests are designed to measure the minimum skills needed to function in high school, it is reasonable to expect all students to pass them. At the same time, other educators are concerned that the tests are not an accurate measurement of the skills of LEP students because of such factors as:

- The length of time required to become proficient in academic English.
- A school system that has lower expectations for LEP students than it does for other students.
- Students' lack of access to content courses.
- Incomplete or interrupted educational backgrounds.
- An environment of poverty.
- Cultural biases in standardized tests.
- Students' lack of familiarity with the testing format.

When districts are under pressure to earn high test scores, there may be a tendency to exempt LEP students from the Basic Standards Tests because it is believed that these students do not have enough English skills to pass the tests. One of the goals of the Minnesota Department of Children, Families and Learning (CFL) and the National Center on Educational Outcomes (NCEO) is to encourage districts to include LEP students in the assessments so that the needs of these students are considered when making future educational policy decisions. Furthermore, including LEP students in the testing will allow educators to determine what impact current language support programs are having and what additional kinds of support are needed (O'Malley & Valdez Pierce, 1994; Rivera & Vincent, 1996; Zehler, Hoptstock, Fleischman & Greniuk, 1994).

In the past, few school districts outside of the metropolitan areas of Minneapolis and St. Paul had substantial populations of LEP students; therefore, educators and policy makers in suburban and rural areas may not have needed to familiarize themselves with the needs of non-native

English speakers. Today that is changing as groups of Southeast Asian refugees resettle in small rural towns and migrant workers arrive for several months during the harvesting season. Suburban and rural areas are seeing the numbers of LEP students grow quickly. There is now a need for all educators and policymakers across the state to become familiar with the needs of LEP students and to learn about the issues involved in assessing them. This document was written for the purpose of providing some important information on the needs of LEP students.

Some of the specific assessment issues that this document will address are:

- The role of testing in education today and the different types of tests mandated by educational reform legislation —*Why have tests become so important?*
- Characteristics of LEP students and relevant demographic data — *Who are LEP students?*
- Factors within the student, the student's environment, and in the testing environment that can influence the test results — *What do you need to be aware of when you choose a test and implement it?*
- Modifications and adaptations that make assessments more appropriate for LEP students — *What support can we give to include more LEP students in assessments and to get an accurate picture of their academic ability?*
- Research findings related to current assessment practices in states and school districts across the country — *What are other states doing and what can we learn from them?*
- Research findings from large scale assessments at the national level, such as the National Assessment of Educational Progress (NAEP) — *What does the research on large scale assessments tell us about the best way to assess LEP students?*

In this document we focus on literature pertaining to best practices in assessing LEP students. It is based on an annotated bibliography of resources related to assessment and LEP students (Liu, Thurlow, Vieburg, El Sawaf & Ruhland, 1996). It does not address the topic of second language acquisition in detail, nor competency development in content areas.

To find material for the annotated bibliography on which these reviews are based, the Psychology Literature database, the ERIC database, and the World Wide Web were searched. Additionally, other materials were obtained from state and national agencies (i.e., Center for Applied Linguistics, North Central Regional Educational Laboratory [NCREL]), as well as from area

libraries and university bookstores. The search focused on documents published within the past 10 years, but it included earlier resources if it was thought that they were still relevant.

Due to the wide variety of topics covered in the annotated documents and articles, only those that are relevant to the topic of high stakes graduation assessments are included in this literature review. Some publications mentioned in this document were not listed in the annotated bibliography because they were obtained after the bibliography was published.

The Important Role of Testing

Why is testing so important in education today? Before turning to a discussion of LEP students and assessment, it is helpful to consider the relationship between testing and the broader educational context. An understanding of this relationship will help frame the discussion of the needs of LEP students.

The Relationship of Educational and Societal Trends to Assessment

Peter Airasian, in his article *State Mandated Testing and Educational Reform: Context and Consequences* (1987), points out that "As new educational roles and expectations emerge from shifting societal priorities and power bases, new testing roles and expectations also emerge to complement altered educational priorities" (p. 398). Testing then, has a direct relationship to the role of education in society, and that role changes as society changes.

According to Airasian, four trends have developed over the past 20 years in American society that have greatly influenced the educational system:

1. Americans now have a greater tendency to diagnose societal problems and to look for ways to solve them.

2. There is a growing belief in America that social problems are related to undesirable behavior in people instead of being related to problems in the social system, and that these undesirable behaviors can be changed. The easiest time to affect behavior is when a person is young.

3. There is a greater push for equity and equal opportunity for underrepresented social groups, especially those that have not been served equally by the educational system (Airasian, 1987; National Coalition of Advocates for Students, as cited in Fradd, McGee, & Wilen, 1994; Willshire Carrera, as cited in Fradd, et

al., 1994). This trend has "focused attention on the results or outcomes of education" (p. 396).

4. American family life has changed so that schools and families no longer have as supportive a relationship as they once had. Because of the changing structure of families, schools have taken on some of the responsibilities that families used to have.

All four of these trends support social reform. Society sees schools as a practical place to locate social reform efforts because there are large groups of young people from diverse backgrounds. The school can implement and have control over programs more easily than other institutions and can reach children at an age when behavior is more easily influenced.

As a result of such societal trends, there have been new educational trends (Airasian, 1987). According to Airasian, first, schools are expected to offer more types of programs to meet the varying needs of a more diverse student population. Because of this expansion, education has become an economic enterprise. Second, state governments have taken more control over education because of the increased costs of education and the larger number of goals that educational institutions must meet. Third, education has become more political as a larger number of societal forces influence decision making.

When the societal and educational trends are combined, assessment takes on a new role. State mandated testing programs are now used to monitor the entire educational system and to give proof of each individual student's performance within that system. The general public tends to see assessments as less biased than educators' judgments of how well schools are doing, so there is public support for testing programs. However, as different social groups bring pressure on the system to serve the needs of their students, the tension between the goals of these different groups becomes more apparent (Airasian, 1987; National Coalition of Advocates for Students, as cited in Fradd et al., 1994; Willshire Carrera, as cited in Fradd et al., 1994). For example, there is tension

> between the use of tests to improve educational quality and current definitions of equal educational opportunity (Astin et al., 1982; Samuda, 1975). Test results used to carry out policies designed to raise academic standards or certify pupil and teacher competence inevitably clash with other policies aimed at equalizing educational opportunity and integrating racial/ethnic groups in schools and classrooms. (Airasian, 1987, p. 407)

These conflicting goals of obtaining high educational quality and of providing equal educational opportunity for all students are at the center of the educational reform legislation that has been put into place over the past 20 to 30 years (Airasian, 1987; Fradd et al., 1994). The goals also

drive the current move toward increased accountability for schools, the increased use of testing as the basis of judging the effectiveness of schools, and increased concern over how minority students, in particular LEP students, are being served by the educational system.

Educational Reform Legislation

Four major pieces of federal legislation have addressed the issue of educational quality:

- *The National Education Goals of 1989* listed educational goals related to improving the quality of education by the year 2000 (Zehler et al., 1994).

- *The Goals 2000-Educate America Act of March, 1994* put the national education goals into law, provided a framework for education reform (August, Hakuta, & Pompa, 1994), and supported the integration and coordination of federal and state programs so that all students would achieve high standards (August, Hakuta, Olguin, & Pompa, 1995). As a part of the reform, "meaningful, challenging and appropriate" assessments are required for all students (National Academy of Education [NAE], 1996).

- *The Elementary and Secondary Education Act (ESEA)* emphasized full opportunities for all students to achieve at high levels (Garcia, as cited in Zehler et al., 1994).

- *Title I of the Improving America's Schools Act (IASA) of 1994* reauthorized the ESEA for five more years and refocused it (Anstrom & Kindler, 1996; Wilde & Sockey, 1995; Wilen & van Maanen Sweeting, 1986; Zehler, et al., 1994). The purpose of this piece of legislation was to ensure that Title 1 students are also challenged with the same high expectations as other students. Programs for educationally disadvantaged students, including limited English proficient (LEP) students, were funded and other measures were taken to ensure high educational achievement for such students. Assessment is also a part of the IASA, and LEP students are required to be included in the testing programs "to the maximum amount practical with reasonable adaptations" (Anstrom & Kindler, 1996, p. 8). The legislation also requires an effort be made to obtain assessments in non-English languages.

Clearly all types of testing (classroom and large-scale district or state) have a prominent role in education today. Because assessment is mandated by laws addressing educational quality, schools need to address issues of equitable assessment for all students now instead of waiting in the

hope that federal mandates will change. Lacelle-Peterson and Rivera (1994, p. 56) state:

> U.S. educators must reflect seriously on the implications of reform, and, in particular, assessment reform, for those students who have not, historically, thrived automatically. Unless we do so, current efforts to improve U.S. education will amount to little more than a reformulation, perhaps with new labels, of the status quo.

One group of students who has not "thrived automatically" is made up of language minority and limited English proficient students. Anstrom and Kindler (1996) highlight the important role students play in educational reform by writing that as the nation works toward meeting the goals in Goals 2000, "our ability to achieve these goals is increasingly dependent upon the school's ability to educate language minority students" (p. 1). If the goal of educational reform is to educate all students to high standards, then it is important to have a comprehensive understanding of the non-native English speaking student population, the challenges these students face, and the recommendations for improving their educational achievement as measured through large-scale assessments (Anstrom & Kindler, 1996, p. 1).

Characteristics of LEP Students

Definitions

Any discussion of issues concerning non-native speakers of English must start with a review of the terminology used. There are many different terms used to describe such students and it is difficult to review the literature on assessment without a standard set of terms. For the purposes of this document, we have chosen to use the following terms:

> **Linguistic/language minority (LM)**. A student who speaks a language other than English at home.

> **Limited English proficient (LEP)**. A student who is a linguistic minority and who, in English, performs at a level below his or her native English speaking peers. The student's English ability makes it difficult for him or her to benefit from English-only instruction. The term "LEP" may be controversial because it focuses on the student's lack of ability in English; however, we have chosen to use this term because the federal government uses it. Our definition of LEP is purposely very broad because there is no standard definition.

According to Cheung and Solomon (as cited in Anstrom, 1996), most states use either an operationalized version that fits the state's situation, or a simplified version of the following federal definition from *Title VII of the Improving America's Schools Act of 1994* (Public Law 103-382) (this definition can also be found in Council of Chief State School Officers [CCSSO], 1992a, 1992b; Delvecchio & Guerrero, 1995; Navarette & Gustkee, 1996): A student is LEP if he or she

> has sufficient difficulty speaking, reading, writing, or understanding the English language and whose difficulties may deny such individual the opportunity to learn successfully in classrooms where the language of instruction is English or to participate fully in our society due to one or more of the following reasons:
>
> - was not born in the United States or whose native language is a language other than English and comes from an environment where a language other than English is dominant;
>
> - is a native American or Alaska native or who is a native resident of the outlying areas and comes from an environment where a language other than English has had significant impact on such individual's level of English language proficiency; or
>
> - is migratory and whose native language is other than English and comes from an environment where a language other than English is dominant. (sec. 7501)

The federal definition is often used to determine eligibility of students for bilingual education services or to design entry and exit criteria for ESL programs (DelVecchio & Guerrero, 1995). Below are definitions of limited English proficiency from three states. These definitions clearly illustrate the variations that exist in the term "LEP":

> **From Texas:** A student may be classified as LEP by the language proficiency assessment committee if one or more of the following criteria are met: (i) the student's ability in English is so limited or the student is so handicapped that assessment procedures cannot be administered, (ii) the student's score or relative degree of achievement on the agency-approved English proficiency test is below the levels established by the agency as indicative of reasonable proficiency, (iii) the student's primary language proficiency score as measured by an agency-approved test is greater than his proficiency in English or, (iv) the language proficiency assessment committee determines, based on other information such as (but not limited to) teacher evaluation, parental viewpoint, or student interview, that the student's primary language proficiency is greater than his proficiency in English or that the student is not reasonably proficient in English. (Texas Education Code s21.455, as cited in Anstrom, 1996)
>
> **From California:** Each student with a home language other than English. . .has been assessed. . .in English comprehension, speaking, reading, and writing. . . .

> Each LEP student has a score of less than fluent according to publisher's norms on a state-authorized test of comprehension and speaking proficiency (K-12), or a score of fluent on the English comprehension and speaking proficiency test AND a score below the district-established standards on the district's English reading and writing assessments. (California Education Code, as cited in Anstrom, 1996)
>
> **From Minnesota:** For the purposes of identifying students for placement in a language program, Minnesota defines LEP as "a pupil in any of the grades of kindergarten through 12 who meets the following requirements: a) The pupil, as declared by parent or guardian 1) first learned a language other than English, 2) comes from a home where the language usually spoken is other than English, or 3) usually speaks a language other than English; and b) the pupil's score is significantly below the average district score for pupils of the same age on a nationally normed English reading or English language arts achievement test. A pupil's score shall be considered significantly below the average district score for pupils of the same age if it is one-third of a standard deviation below that average score." (Minnesota Statutes 126.262)

It is clear that while these three states include some reference to language proficiency and language background, each state lists different types of characteristics in defining "limited English proficient." Table 1 shows the differences between the definitions in Texas, California, and Minnesota.

On a broader scale, a recent survey of State Education Agencies (Hopstock & Bucaro, as cited in Anstrom, 1996), examined definitions of LEP in 52 states and territories and found the same lack of consistency in them. All 52 states and territories had a definition, but each included different components:

- 44 of 52 specified non-English background.
- 29 of 52 specified difficulty speaking, reading, or writing English.
- 17 of 52 indicated a percentile cutoff on a language or achievement test.
- 9 of 52 allowed for local determination of LEP status using local criteria.
- 13 of 52 used other criteria (e.g., grades, teacher judgment).

The major issue at the heart of this disagreement over the term LEP is a lack of agreement on what the term "fully English proficient" means, how to assess a student's English proficiency, and which modes of English (listening, speaking, reading, or writing) to assess. For more information on assessing language proficiency see Delvecchio and Guerrero (1995).

Table 1. *Similarities and Differences in Definitions of LEP in Three States*

Characteristics Listed in Definition	States		
	Texas	California	Minnesota
Test Data			
Score below agency established cut score	X		
Score below average district score			X
Score below publishers norms on commercial test		X	
Low score on state authorized language proficiency test		X	
L2[1] proficiency is adequate but low score on reading & writing test		X	
Must take agency approved L2 test	X		
Must take agency approved L1 test	X		
Must take nationally normed reading & language arts test in L2			X
Gives specific cut off score on test			X
Language Proficiency			
Unable to be tested on L2 test because of low L2 proficiency	X		
Must be tested in 4 modalities (reading, writing, speaking, listening) in L2		X	
Shows higher proficiency on an L1 test than an L2 test	X		
Language Use			
First language is not English			X
Most frequently used language is not English			X
Home language is not English		X	X
Who is involved			
Parent must be involved		X	
Parent can be involved			X
Committee determines LEP status	X		

[1] L1 = First language.
L2 = Second language.
Note: Some LEP students may speak English as their 3rd or 4th language, but in this case we use L2 to refer to English.

BEST COPY AVAILABLE

Variations in definitions and in the way in which districts apply them may arise for several other reasons as well:

- Definitions are often vague and may not specify the test to be used or a cut-off score. In many states, local education agencies can choose which language proficiency test and which cut-off scores to use. A student tested in two different states with two different tests could be placed in a language assistance program in one state and not in the other (CCSSO, 1992b). An additional characteristic of vague definitions is that they can allow more students to be excluded from assessments and thereby improve overall district and state scores (Zlatos, 1994).

- States may have different definitions of LEP for different purposes. For example, the state of Minnesota has one definition of LEP to identify students in need of services and another, much broader definition of LEP for the purpose of determining assessibility on statewide tests: *Definition in Rule: According to Minnesota Rule 3501.0100, LEP students are "individuals whose first language is not English and whose test performance may be negatively impacted by lack of English language proficiency."*

- States generally recommend identification and program placement procedures for LEP students, but most do not require districts to use the recommended procedures (O'Malley & Valdez Pierce, 1994).

- ESL and bilingual education programs may have needs that conflict with accurate labeling of students (Gandara & Merino, 1993). For example, a program may be overcrowded and advanced LEP students may be prematurely exited to make space for students with lower skills (O'Malley & Valdez Pierce, 1994). The exited students are no longer classified as LEP but they may still have the same needs for bilingual or ESL services.

Effects of Varying Policies and Definitions

Varying policies among states in the application of the term LEP can have several effects (Gandara & Merino, 1993). One major effect is that language minority students may be denied access to programs that would help them achieve at levels comparable to their native speaking peers because of the process that schools follow in the identification and placement of LEP students (O'Malley & Valdez Pierce, 1994). A lack of access to language learning programs influences the students' opportunity to learn academic content. A second major effect is that estimates of the number of LEP students, the location of these students, and the number of students being served by a language learning program vary greatly (CCSSO, 1990). A lack of precise information

prevents educators and policymakers from providing effective educational programs for LEP students (CCSSO, 1990; Fradd et al., 1994; Gandara & Merino, 1993).

In response to the dilemma created by different definitions of the term LEP, many educators and researchers have called for an objective, standardized definition that would be used throughout the country (Gandara & Merino, 1993). A standardized definition of LEP would have both positive and negative aspects. On the positive side (Anstrom, 1996) a standardized definition would lead to:

- A better understanding of LEP students.
- A more precise estimate of the LEP population.
- A more accurate estimate of how that population is distributed across the country, allowing for a fairer distribution of resources.
- The use of a common term for discussing students.
- The assurance that a student identified as LEP in one state would also be identified as LEP in another state if he or she moved.

On the negative side, a standardized definition of LEP could also be:

- Too explicit or too broad, resulting in an incorrect number of students identified as LEP.
- Favor one subset of the language minority population over another one.
- Increase the cost and the hardship for districts because more testing is required. (Hopstock & Bucaro, as cited in Anstrom, 1996)

Variability in definitions and policies necessarily creates problems in estimating the size of the population needing services. This, in turn, creates a number of problems in providing those services.

Estimated Population Size and Educational Status

It is important to have accurate information about the size and make up of the LEP population in order to plan effective programs (Fradd et al., 1994). Estimates of the size of the language minority and LEP population in the United States clearly indicate the difficulty involved in measuring the population (Anstrom, 1996). There are two ways that researchers measure the

language minority and LEP populations. First, many of them rely on the latest census administration for data about the size and ages of the linguistic minority population. According to the 1990 census, about 16% of the 91.7 million households at that time were language minority households, and about 50% of them had school-aged children (Waggoner, as cited in Zehler et al., 1994; Zehler et al., 1994). This source of information reported that the total number of language minority students in grades K-12 was about 9.9 million out of a total population of 44.9 million students (Anstrom, 1996).

Though these numbers appear to be straightforward and specific, they may be inaccurate for several reasons: (1) the census relies on self-reported data and this type of data may be inaccurate (Anstrom, 1996); (2) the census does not accurately represent all of the linguistic and cultural sub-populations in the country because it only asks for information about the 15 most common sub-populations (United States General Accounting Office [USGAO], 1994); and (3) the census relies on the ability of the individual to fill out a fairly complex form; therefore, certain members of the language minority population (in particular those with low literacy levels in English) may not have been able to provide data.

The second way of obtaining information about the language minority and LEP student population is through school-based studies (Anstrom, 1996). Though these studies have more sophisticated ways of obtaining data, the data again vary greatly due to the inconsistent definitions of LEP across states. School-based studies estimate the LEP student population at anywhere from 1.3 million to 7.5 million, depending on the method chosen to measure limited English proficiency and the level of proficiency used as the standard (Anstrom, 1996). The large variation in numbers also is related to district reporting practices. Some districts report all students who are eligible for LEP services, while others report only those receiving services (CCSSO, 1990). The commonly accepted number of LEP students in the U.S. is about 3.5 million (CCSSO, 1990). Approximately 73% of LEP students come from Spanish-speaking backgrounds (CCSSO, 1990). Each of the following language groups represents approximately 4% of the student population: Vietnamese, Hmong, Cambodian, Cantonese, and Korean. The smallest group, at 2-1/2 %, is made up of students who speak a Native American language (Fleischman & Hopstock, as cited in Anstrom & Kindler, 1996; Navarette & Gustkee, 1996).

Both the census and the school-based data point to the rapid growth of the linguistic minority population in recent years, and LaFontaine (as cited in CCSSO, 1990) states that this population is expected to grow about 32% by the year 2000, with the subgroup of LEP students growing about 17%. For more about demographic changes in the U.S. population, see Hodgkinson (1992) and Hodgkinson and Outtz (1992). The number of students needing language support programs is certain to increase. It is important to look at the number of students being served by language support programs now so that the future increased needs can be planned for and met.

According to school-based data, LEP students could be 3–17% of the school population, depending on which estimates are used (Anstrom, 1996). In the 1992-93 school year, more than 2.4 million LEP students were receiving educational support in special programs. Students were spread out among several programs: Chapter 1 served 31%, Special Education served 6%, Chapter 1 Migrant Education served 10%, and Title VII Bilingual Education served 11% of the students. Twenty-one percent of LEP students were not served by any special program (Anstrom & Kindler, 1996). Many LEP students receive insufficient English language instruction and little or no first language instruction (Olson, as cited in Cuevas, 1996). A 1987 survey of SEAs supports the finding that a large percentage of LEP students were not receiving any LEP or bilingual services at that time (CCSSO, 1990). The survey results showed that 20 of 32 states did not provide at least 25% of the LEP student population with any language support services, 4 of 32 did not provide services for at least 60%, and 1 of 32 states did not provide services for any LEP students. Twelve of the 32 states reported that all the students identified as LEP received services, but the researchers found that barriers like limited funding and insufficient data prohibited schools from actually serving all of the LEP students in those states.

Other Characteristics of LEP Students

In general, the majority of LEP students are young, Spanish-speaking people (Cuevas, 1996; Navarette & Gustkee, 1996). Statistics show that more than 65% of these students are currently in elementary school while only 14% of them attend grades 10 through 12 (Navarette & Gustkee, 1996). In spite of LEP students' common difficulties with English, they have a variety of diverse backgrounds that affect their educational needs; these students should not be seen as a homogenous group (Lacelle-Peterson & Rivera, 1994). If educators expect LEP students to succeed academically, they must recognize this diversity instead of treating LEP students alike. As an entire group, LEP students have the potential to make a substantial contribution to the American economy by the time they are old enough to work (National Commission on Testing & Public Policy, as cited in Fradd et al., 1994). This contribution may be a positive one if students have received a solid education; or it may be a liability if people with limited English ability lack a high school diploma or have few job skills and need government assistance to survive, or are spending time in prison. According to demographic research, there is a strong relationship between violent crime rates and high school drop out rates (Hodgkinson & Outtz, 1992). Approximately 82% of prison inmates are high school dropouts (Hodgkinson, 1992).

Canales (as cited in Gandara & Merino, 1993) states that in addition to language difficulties, many LEP students possess characteristics associated with low educational achievement, particularly with low test scores. Such characteristics include:

- *A tiered education system* in which large numbers of LEP students are exposed

to a different curriculum with lower standards and therefore do not graduate with the "rigorous preparation needed to compete in the job market" (August et al., 1994, p. 14). In a study of California schools, Minicucci and Olson (1992) found that approximately 48% of schools in the state offered few to no content classes for LEP students. These findings may hold true for other states as well.

- *Difficulties adjusting to a new school culture*, which may be a result of different teaching methods, different expectations for student behavior in class, different daily routines, and different relationships between teachers and students. (Minicucci & Olson, 1992).

- *Low social status* because of an increase in anti-immigrant feelings and racial tensions in schools (California Department of Justice, as cited in Minicucci & Olson, 1992).

- *High family mobility* resulting from immigration status and migratory work that creates a lack of educational continuity (Anstrom & Kindler, 1996; CCSSO, 1990; Minicucci & Olson, 1992; USGAO, 1994).

- *Poor academic preparation* due to limited resources of schools in lower income neighborhoods, or to LEP students' lack of access to content courses (Canales, as cited in Gandara & Merino, 1993; Minicucci & Olson, 1992). "More than 40% of all LEP students live in large urban districts where resources to educate these students are often scarce" (Anstrom & Kindler, 1996).

- *Poor attendance* reflected by the fact that 20% of language minority students have missed at least two years of high school, and 12% have missed at least two years of junior high (Fleischman et al., as cited in Anstrom & Kindler, 1996)

- *Poverty* and related factors such as high unemployment rates, substandard housing, overcrowded neighborhoods, and health problems that interfere with a student's ability to learn (CCSSO, 1990; Prewitt Diaz et al., 1989, as cited in Anstrom & Kindler, 1996; United States General Accounting Office, 1994). According to Hodgkinson and Outtz (1992), "The number one item which predicts the percent of children who will be at risk of school and health failure is poverty" (p. 5).

- *A higher drop out rate* because of poverty (Hodgkinson & Outtz, 1992; Pallas, as cited in McDill, Natriello & Pallas, 1985), linguistic and cultural differences that act as a barrier to educational achievement (Minicucci & Olson, 1992), poor grades, poor relationships with teachers (McDill, et al., 1985), and more

pressure on students to work and help support their families (Chavkin, as cited in Anstrom & Kindler, 1996; Minicucci & Olson, 1992; Prewitt Diaz et al., as cited in Anstrom & Kindler, 1996). Immigrant families are more than twice as likely to be poor (U.S. Census 1990, as cited in USGAO, 1994). Based on 1990 census data, it is estimated that 37% of language minority students live in poverty, in contrast to 17% of the total student population (USGAO, 1994). In general, states with a high percentage of children living in poverty tend to have a higher dropout rate and a higher rate of teen pregnancies (Hodgkinson & Outtz, 1992).

Migrant, Immigrant, and Refugee Students

Migrant Students. Migrant students are one group with special needs that make up a portion of the LEP population. These students often have a greater need for LEP services than estimates of the LEP population reflect, and traditional programs designed to meet the needs of a stable group of LEP students may not succeed with migrant students (Anstrom & Kindler, 1996).

Approximately 80% of migrant students are Spanish speakers (Henderson et al., as cited in Anstrom & Kindler, 1996), and a majority of them may come from homes where adults speak little or no English. Two-thirds of migrant families have an income below the poverty level (Strang et al., as cited in Anstrom & Kindler, 1996), as a result, students may experience disabling poverty-related health problems that can affect their ability to attend school and to learn (Huang, as cited in Anstrom & Kindler, 1996). In addition, because migration often is based on the agricultural seasons instead of on the school year, a student's educational background may have many gaps (Anstrom & Kindler, 1996). Students may be exposed to different curricula and theories of education in each school they attend, credits may not transfer from one school to another, and educational records may be incomplete because information does not get passed from one school to the next (Cox et al., as cited in Anstrom & Kindler, 1996). Migrant students may be identified as LEP in one place and not in another, or they may be eligible for language support services and never receive them. For these reasons, it is often difficult for migrant students to stay at grade level and to achieve academically (Salerno, as cited in Anstrom & Kindler, 1996). As a result, the migrant student population tends to have a high dropout rate and a low graduation rate (Anstrom & Kindler, 1996).

Immigrant and Refugee Students. Immigrant and refugee students may be subdivided into three separate categories of students (Minicucci & Olson, 1992):

- **Immigrant students with a good education from schools in their native countries.** These students may have some initial difficulties in adjusting to a new culture and a new educational system, but they generally are able to overcome the difficulties. The longer students stayed in school in their native countries, the higher their educational achievement tends to be in the second culture (Saville-Troike, 1991). In part, higher educational achievement is due to the transfer of certain school-oriented behaviors and knowledge about the types of experiences a student typically has in school (Saville-Troike, 1991; Saville-Troike & Kleifgen, as cited in Adamson, 1993).

- **Students who have been in American schools and have participated in language programs since elementary school**. These students may complete a kindergarten through 12th grade ESL program and still have insufficient English skills to be mainstreamed with native speaking peers.

- **Students with little prior schooling and minimal literacy skills**. These students may come with significant health and emotional needs caused by war in their native countries and time spent in refugee camps (Anstrom & Kindler, 1996; USGAO, 1994). The emotional needs must often be met before academic achievement in English can take place. Students' education may have been interrupted by conditions in their native countries, they may have low literacy levels in their native language, their knowledge base may be limited to the traumatic experiences they have had in life, and they may have had little or no opportunity to develop the knowledge frameworks they need to succeed in an academic setting (Chamot, 1992).

As a result of low literacy levels in the native language, low levels of English proficiency, incomplete educational backgrounds, and significant emotional needs that act as barriers to education, immigrant and refugee students in the last two subgroups may be more likely to be retained a grade or put in a track with a less rigorous academic curriculum (CCSSO, 1990; NCAS, as cited in Cuevas, 1996).

Factors that Affect Test Results for LEP and Language Minority Students

It is easy to assume that once LEP students learn English, they have the same testing needs as students who are native speakers of English and can perform equally well on the same tests. However, research suggests that non-native speakers of English have unique needs because of such factors as the influence of their native language and culture, bilingualism, and social status. In addition, factors within the testing situation can have a negative impact on the scores an LEP or language minority student obtains, and reduce the validity and reliability of the test for these students (Duran, 1989). If a test administrator or evaluator is not informed about these factors and their impact on the assessment process, the results of the test may be misinterpreted (Zehler et al., 1994). Once misinterpreted data are in a student's file, they can be used to make high stakes educational decisions without any consideration being given to the inequities of the testing situation (Lam, 1993; Zehler et al., 1994). The factors mentioned above must be taken into account in a testing situation (Chamberlain & Medeiros-Landurand, 1991) so that they do not lower test scores and prevent the tests from accurately predicting the educational success of both the student and the school (Lam, 1993; Lam & Gordon, 1992; Zehler et al., 1994).

Acculturation and Language Proficiency

The first set of factors comes from within the student and includes the student's level of acculturation and proficiency in English. It is important to consider the student's level of acculturation to the United States (Chamberlain & Medeiros-Landurand, 1991; Cloud, 1991; Olmedo, 1981). According to Olmedo (1981), acculturation involves "the acquisition of [the] language, values, customs and cognitive styles" of the new culture. This factor is not only important for students who were born in another country, but also for students who have grown up in an ethnic community in the U.S. that has a distinct language and culture that is different from the mainstream. Not all LEP students were born outside of the United States (Navarette & Gustkee, 1996). According to elementary schools, 41% of language minority and LEP students were actually born in the United States.

"A student who is totally unacculturated to the mainstream culture is clearly at a disadvantage in the assessment process" (Chamberlain & Medeiros-Landurand, 1991, p. 132) because the student may respond to questions differently from his or her mainstream peers and may not have similar background knowledge and experiences. These students may also be at a disadvantage in the testing process because they possess a different set of cultural values and beliefs, and these beliefs affect their behavior (Bracken & Barona, 1991; Chamberlain & Medeiros-Landurand, 1991). For example, Deyhle (as cited in Chamberlain & Medeiros-Landurand, 1991) cites research showing that Navajo second graders and their Anglo-American

classmates had different cultural beliefs about testing. The Anglo-American students saw testing as a way to prove their individual abilities and to demonstrate what they could achieve through competing with their classmates. These students were very motivated by tests. In contrast, the Navajo students did not value competition and did not see the tests as being different from other types of classroom activities, so they were not motivated by the testing situation.

According to Chamberlain and Medeiros-Landurand (1991), there are many cultural variables that may influence test results if students are not completely acculturated to the United States (for more on cultural variables see also Greenfield, as cited in Fradd et al., 1994; Westby & Rouse, as cited in Fradd et al., 1994) including:

- *Attitudes toward competition.* Standardized testing in the U.S. is based on the belief that it is good to compete against your classmates and to perform better than they do in order to demonstrate your ability (Goldman & McDermott, as cited in Chamberlain & Medeiros-Landurand, 1991). Not all cultures share this belief. In some cultures, it is important not to hurt other members of your group or make them feel ashamed by performing better than they do. Students from a culture that values cooperation over competition may be at a serious disadvantage in a standardized testing situation (Seymour, as cited in Chamberlain & Medeiros-Landurand, 1991).

- *Attitudes toward the importance of the individual versus the importance of the group or family.* In mainstream American culture, students are taught to value their role as individuals, therefore the testing situation motivates them to perform their best. In other cultures, this same type of motivation may not exist if a person's role as a group member or family member is seen as more important than his or her role as an individual. In such a case, a student must be given a "motivational rationale" that fits with his or her value system (Chamberlain & Medeiros-Landurand, 1991). For example, a student who places more importance on his or her family role can be encouraged to do well on the test so the family will be proud of the student's achievement (Gallimore, as cited in Chamberlain & Medeiros-Landurand, 1991).

- *A belief in fate versus belief in individual responsibility.* Americans tend to believe that a student has some control over the testing situation and that he or she is therefore partially responsible for the outcome of the test. In contrast, students from other cultures may believe that fate, God, or many gods control their life and the results of their test. Students with this type of cultural belief may not be motivated to perform well or to improve their performance because they may feel that the situation is out of their hands.

- *Gender roles.* A study by the National Center for Education Statistics (1992) has found that the most important factor in academic performance is the parents' expectations that a student will graduate from high school and go on to college. In mainstream American culture, this expectation may apply equally to boys and girls, but in other cultures it may not. Educational achievement may not be valued for girls in some cultures, or in some families within a cultural group, and this lack of value can affect the test performance of the female students if expectations are lower for them than for male students. In addition, Gollnick and Chinn (as cited in Chamberlain & Medeiros-Landurand, 1991) suggest that some male students may perform better with male test administrators because the student has a "high regard" for the male as a role model.

- *Attitudes toward the use of time.* Standardized testing is based on the belief that the speed at which you finish the test demonstrates your skill (Kurt, as cited in Chamberlain & Medeiros-Landurand, 1991). In addition, the test usually only allows students to work on one section at a time and to only work on the test. Activities such as talking to one's neighbors, or working on more than one section at a time, are not valued in the United States. However, some cultures do not share this belief that it is good to work on only one thing at a time and to finish it as quickly as possible. Students from these cultures may have been taught that the quality of an answer to a question is more important than the time it takes to give the answer, and therefore may be at a serious disadvantage if they are not accustomed to having to finish a standardized test in an allotted time, or do not see the reason for it.

- *Attitudes toward the demonstration of knowledge.* In highly individualistic cultures such as mainstream American culture, children are frequently encouraged to demonstrate knowledge that is obvious or that is already known by others (Fradd et al., 1994; Taylor & Lee, 1987). Tests may require students to choose obvious answers to reading comprehension questions or to display their knowledge in the written expression section. Students from cultures that are more group-oriented, however, "may be reticent to talk about already known topics, and may feel as if they are showing off when explaining something obvious" (Fradd et al., 1994, p. 72).

- *Verbal communication norms.* In a standardized testing situation, students are expected to ask questions when they do not understand the testing directions. However, in some cultures young people are not encouraged to initiate conversation with adults (Fradd et al., 1994; Birdwhistell, as cited in Chamberlain & Medeiros-Landurand, 1991; Crago, as cited in Chamberlain & Medeiros-

Landurand, 1991); as a result, a student from such a culture may not ask questions of the test administrator even when he or she does not understand (Wilen & van Maanen Sweeting, 1986). Additionally, in some cultures asking questions of the teacher implies that the teacher is at fault for not explaining the task clearly enough, and so the student stays silent out of respect for the teacher (see also Bracken & Barona, 1991).

- *Use of body movements and gestures.* In order to be successful in school, and especially on a standardized test, a student must be comfortable with a very passive style of behavior (Almanza & Mosley, as cited in Chamberlain & Medeiros-Landurand, 1991). The test rewards students who can sit quietly for long periods of time, stay in their seats, and focus on only the teacher's directions or the test itself. Students who come from a culture that values a more active style of talking and moving about in the classroom may not be able to focus on the test for a long period of time and will be at a disadvantage. In addition, some gestures may be used differently in different cultures. For example, an American test administrator may nod his or her head to give permission when a student asks to use the bathroom during a testing situation; however, in the student's culture that same gesture may mean that permission is denied. A student who misinterprets this gesture may have to finish the test in discomfort and the test score may be negatively impacted as a result.

- *Proximity*. In a standardized testing situation, students are usually separated so that they cannot see anyone else's paper or disturb other students by touching them. This enforced space between students may feel very uncomfortable to people from cultures that value a much closer proximity to and more physical contact with people around them. This discomfort may add another dimension of stress for the student from another culture.

- *Use of eye contact.* In a standardized testing situation, eye contact with the test administrator becomes very important when the administrator is giving directions or when the testing has begun and oral communication is not encouraged. However, in some cultures it is not polite to look teachers in the eye because teachers are in a position of respect. A student who looks at the floor while the administrator is giving test directions may miss vital parts of the directions, which in turn may affect overall test performance.

The existence of all of the cultural variables mentioned above points to the need for educators, and especially test administrators, to be sensitive to the hidden obstacles that students may face in the testing situation. However, Bracken and Barona (1991) caution educators to keep in mind

that individual differences may play more of a role than culture for some students. For example, a child of Vietnamese background who was raised in France may have needs that differ from a Vietnamese child who grew up in Vietnam. Ideally, for each LEP and language minority student, educators should identify individual and cultural variables that may affect the results of the particular assessment the student is taking (Chamberlain & Medeiros-Landurand, 1991). Furthermore, the test itself should be examined to see whether there is any specific cultural content that may be problematic for these students from other cultural backgrounds, and whether the cognitive styles required on the test are styles that the students are not accustomed to using. Consideration of these factors will help educators account for differences between a student's knowledge and performance on the test (Chamberlain & Medeiros-Landurand, 1991).

Another intrinsic factor that must be considered in the testing situation is a student's second language proficiency. The idea of language proficiency is crucial to any discussion on large scale testing because any test written in English is really a test of English proficiency as well as a test of content knowledge (August et al., 1994; Lacelle-Peterson & Rivera, 1994; National Council on Measurement in Education, as cited in Lam, 1993; Navarette & Gustkee, 1996). Writers of standardized tests usually assume that test takers have no language difficulties that interfere with test performance (Lam & Gordon, 1992), and the writers do not take into account the role of second language proficiency (Zehler et al., 1994).

Federal law requires language proficiency testing to identify students for ESL services. The results of these proficiency tests in turn affect which students are eligible to receive certain types of accommodations on large scale assessments and the types of educational opportunities provided to these students. While most educators recognize a student who is "fully English Proficient" (FEP) (see CCSSO, 1992 for a complete definition of FEP), there is no common understanding of the nature of proficiency and how best to measure and describe the proficiency of students who are not fully English proficient (Delvecchio & Guerrero, 1995; Guerrero & DelVecchio, 1996; Oller & Damico, 1991). This lack of agreement on the nature of language proficiency plays a major role in the difficulties involved in identifying limited English proficient students for placement in an ESL program or for exiting from an ESL program (Gandara & Merino, 1993). Several controversial theories on the nature of language proficiency exist (Delvecchio & Guerrero, 1995). The theory an educator ascribes to generally influences the choice of a type of proficiency test; subsequently, the test influences which students are determined to be "limited English Proficient" in a particular state, school, or district (Ulibarri, Spencer & Rivas, as cited in DelVecchio & Guerrero, 1995). Here are four ways of viewing proficiency in a language:

1. **Structuralist Theories** (DelVecchio & Guerrero, 1995; Guerrero & DelVecchio, 1995; Oller & Damico, 1991). The theories within this category describe language proficiency as a conglomerate made up of separable bits and pieces of knowledge

about the structure of a language (i.e., the sounds of the language, the vocabulary, the grammar, etc.) and also of specific language abilities (i.e., speaking, listening, reading, writing). To be proficient in a language, a person must learn each skill so that he or she possesses a large repertoire of well learned habits on which to draw. If an educator is assessing an LEP student's proficiency based on this type of theory, he or she would choose a "discrete point" test that does not involve knowledge of content and which allows him or her to assess each component of language in isolation. Then he or she would combine the results from each component to form the total picture of a person's proficiency. Test developers tend to use discrete point testing most often (DelVecchio & Guerrero, 1995). The Test of English as a Foreign Language (TOEFL) is an example of a "discrete point" test (Adamson, 1993).

2. **Common Factor Theories** (DelVecchio & Guerrero, 1995; Guerrero & DelVecchio, 1996; Oller & Damico, 1991). The theories in this category view language proficiency as a conglomerate of individual parts, similar to the discrete point theory. However, proficiency is seen as more than just knowledge of language structures; it also involves knowledge of phonology, syntax patterns, and vocabulary, as well as involving background knowledge and reasoning ability. In Common Factor theories, a proficient person uses more than one component of language at a time (listening, speaking, reading, and writing), so he or she should be able to, for example, listen to something and respond to it orally. If you are testing an LEP student's proficiency based on this type of a theory you might ask the student to listen to something and then speak about it, or read something and then write about it.

3. **Communicative Competence Theories** (Adamson, 1993; DelVecchio & Guerrero, 1995; Fradd et al., 1994 ; Guerrero & DelVecchio, 1996). Again, these theories describe language proficiency as a conglomerate of many parts, but the parts are integrated. Additionally, proficiency occurs in real life situations that demand a particular response to a particular situation. A proficient person not only knows the structure of the language but can use it to accomplish something (i.e., apologizing, asking for directions, etc.). If an educator wants to assess an LEP student's proficiency based on this theory, he or she would ask the student to perform a task under the same conditions that the student would perform it in real life. For example, if a student is asked to read a story and write a response to it, the educator would choose a piece of "real life" reading material and would make certain that it is something people actually respond to in writing when they encounter it in real life. Performance assessments and other alternative

systems of assessment such as portfolios, can be assessments based on these theories.

Associated with this more pragmatic view of language learning are theories that there are different types of language proficiencies, which occur at different rates (Cummins, as cited in Hamayan & Damico, 1991; Cummins and Swain, as cited in Hamayan & Damico, 1991; Donaldson, as cited in Fradd et al., 1994). The work of Cummins and of Cummins and Swain suggests that these two types of proficiencies are called "Basic Inter-Personal Communication Skills" (BICS), and "Cognitive Academic Language Proficiency" (CALP). A student learning English as a second language can pick up the BICS by being around native speakers of the language and interacting with them in a natural setting. This type of language occurs in situations where there is a great deal of contextual support (e.g., pictures, visuals, and concrete objects that are in the natural environment) and the situation is not very cognitively demanding. BICS may develop within a few years of coming to the United States and because a student can communicate relatively fluently, he or she appears to be "proficient" in English. However, as the researchers point out, this type of proficiency does not guarantee that the same student has the required CALP to succeed in school. Academic language tends to be "context reduced" (where a reader interacts with a text that is devoid of visual support such as pictures and diagrams) and cognitively demanding. CALP requires a high level of cognitive skill in order to read academic texts, analyze them, synthesize them, and create academically sound pieces of argumentative writing. CALP, then, includes reasoning ability as well as proficiency.

According to the researchers, there is a threshold of BICS that must be achieved in order for CALP to develop. LEP students who come from a non-literate home and who have not developed strong language skills (especially reading skills) in their first language, have never learned how to perform academic reasoning tasks that require more sophisticated cognitive skills. If these students are placed in a mostly English-speaking environment, they may not continue to develop BICS in the first language and may take a long time to develop (or may not develop) CALP in either language.

Under the best circumstances (i.e., the student is literate in the first language, the student had some schooling in his or her native culture, the student comes from a literate home, and is receiving language support services at school), CALP may take anywhere from six to eight years to develop in an LEP student. Once attained, CALP can be transferred from the first language into the second

language. In cases where LEP students are not literate in their first language, do not have CALP, come from homes without native language literacy, and have started school in the U.S. either during the teenage years or as a very young child (young children have most likely not had strong development in their native language), the process of developing CALP can take 10 years or more (Collier, 1987, 1989). There is some evidence to suggest that the CALP needed for math and language arts may take less time to develop than that needed for content areas such as academic reading, social studies, and science (Collier, 1987, 1989). The theory of BICS and CALP might then explain the difficulties many high school-aged LEP students are encountering. These students may be under a great deal of pressure because they are expected to learn academic content at the same time that they are learning English, struggling with cognitively demanding tasks for which they may not be prepared, and dealing with the often painful acculturation process. When LEP students are required to participate in large scale assessments in order to graduate, they may not have enough time to achieve the academic language proficiency and the associated cognitive skills necessary to pass the assessment, particularly in the reading portion (Collier, 1989; Collier & Thomas, as cited in O'Malley & Valdez Pierce, 1994).

4. **The Preference Model Theory** (Spolsky, as cited in Adamson, 1993). This fourth view of language proficiency builds on the Communicative Competence theories. Proficiency is again seen as a conglomerate of individual English skills that are interrelated. However, the model proposes that there are different kinds and degrees of proficiency. Some people, especially people who learned English in an academic setting, may learn all four modalities (speaking, listening, reading, and writing), while some people may learn to understand everything they hear in the new language without ever being able to speak it. Both types of people are considered proficient, but in different ways and in different types of language. According to this theory, knowledge of general vocabulary, linguistic structures, and language functions are part of proficiency, but reasoning and background knowledge are not.

Factors That Affect Second Language Learning

None of the theories state that all students will eventually become proficient in English. The majority of people in the world do become proficient in their first language (Hamayan & Damico, 1991). In a second language, some people become proficient relatively easily, while some people never do, even with a great deal of effort (Hamayan & Damico, 1991). There are two types of

characteristics that contribute to a lack of second language proficiency: (1) factors within the learner, and (2) factors within the learner's environment.

Factors Within the Learner. The first factor related to the learner is cognitive skills. The paragraphs above describe how cognitive functioning may be limited if students have not had a chance to develop those skills in their first language. Another aspect of cognitive skills that needs to be considered is a student's style of learning and the way in which the student processes information (Hamayan & Damico, 1991). If a student prefers to learn English by communicating in natural situations (i.e., role-playing how to order in a restaurant and then going to a restaurant to do it), and the teacher prefers to teach English by showing how to analyze grammar and develop rules, the student's ability to learn may be hampered by the fact that the information is not presented in the way that he or she needs it.

The second factor related to the learner involves the learner's attitudes, feelings and personality (Hamayan & Damico, 1991; Wilen & van Maanen Sweeting, 1986). The theory is controversial, but some researchers (see Hamayan & Damico, 1991, p. 49, for a list) believe that if the learner has positive attitudes toward himself or herself, toward English speakers and toward members of his or her own native language group, then the likelihood of becoming proficient may be increased. If the learner has negative feelings toward any of these three, then the likelihood of becoming proficient in English may be decreased. Furthermore, students who have a relaxed, adventurous, and outgoing personality and are not afraid to take risks with English may tend to become more proficient. However, Hamayan and Damico (1991) caution that this proficiency appears to be more in the areas of speaking and writing than in academic areas. Naiman, Frohlich and Stern (as cited in Hamayan & Damico, 1991) say that one should not assume extroverts are always "good" language learners.

The third factor related to the learner is the learner's proficiency in his or her first language (Hamayan & Damico, 1991). As mentioned earlier, if the student did not go to school in his or her first language and only has BICS in that language, it is much more difficult for him or her to obtain CALP in the second language. August et al. (1995) suggest that the content and performance standards being drawn up by the American Council on the Teaching of Foreign Languages (ACTFL) may be useful in determining a student's native language proficiency.

Factors Within the Learner's Environment. There are two factors within the student's environment that may affect the ultimate attainment of English proficiency. First, the parental and community attitudes about English and about English speakers can have a major influence (Hamayan & Damico, 1991). For example, if an LEP student's community values the use of the native language in daily interaction, then that student is more likely not to attain complete proficiency in English because he or she identifies strongly with the native language (Taylor, as cited in Hamayan & Damico, 1991). Similarly, if the parent or community views English speakers

negatively and does not like to interact with them, this attitude may affect the child's proficiency in English.

The second external factor is home literacy in either the first or the second language (Hamayan & Damico, 1991). Collier (1987, 1989) states that native language literacy is one component of cognitive development that affects the acquisition of English proficiency. If the student grows up in an environment where he or she develops strong native language literacy skills and other skills associated with reading and school, that child may attain academic language proficiency more easily. The academic proficiency in English may come more easily to this child because the reading skills attained in the first language are similar in some ways to the skills needed to read in English. Several researchers believe that reading skills are transferable from one language to another language (for a discussion of this point, see Devine, 1988; Garcia, 1991; Saville-Troike 1991), so first language literacy is important for second language literacy. It is important to remember, however, that literacy is not the only component of cognitive development involved in attaining academic proficiency. When parents interact with a child in the native language and use it to communicate the native culture and beliefs, other types of cognitive development take place that can aid in the development of academic proficiency in English and offset a lack of native language literacy. Hmong students in the United States are an excellent example of this type of situation. Despite the fact that many parents and young people are not literate in Hmong, many Hmong students have attained high academic achievement. This may be due in part to the fact that Hmong elders value their native culture and try to transmit it to the young people through the Hmong language. The dialogue in Hmong that takes place around cultural traditions (e.g., extended family meetings, religious ceremonies, etc.) aids in the students' cognitive development.

Issues of Context and Bias in the Testing Situation

Another set of factors that must be considered is the context surrounding the administration of the test, as well as the test itself. In order to ensure that an assessment is valid for an LEP student three particular aspects of the surrounding context must be analyzed (Cline, 1993). According to Bracken and Barona (1991) and Cline (1993), the first aspect to which special attention must be paid is the social context of the assessment. It is important to consider why a student is learning English as a second language and whether there is a great deal of social pressure placed on the student's ethnic group. A foreign exchange student from a European country who chooses to spend time in the U.S. studying English and attending high school is in a much different position than an immigrant student who must learn English in order to get a job and to survive in a new country. This is especially true at a time when public opinion of immigrants is particularly low. The foreign exchange student and the immigrant student may have similar levels of English proficiency and yet achieve differently on the test for these reasons.

The second aspect of context to be considered is the intellectual context (Bracken & Barona 1991; Cline, 1993). There are certain types of assumptions that educators may make about bilingualism and about students with minority and immigrant status that may affect how the educator interprets test results. In addition to the assumptions that an educator makes, the educator's knowledge of second language acquisition may also have an effect (Cline, 1993). For example, if a student turns in the written portion of a test with code switching — the use of two languages in the same writing sample — a test administrator with little knowledge of bilingualism may react negatively and believe that the student's English is less than fluent because both languages are present. On the other hand, a test administrator who knows something about bilingualism may look at the same sample and understand that the student mixed the two languages for a specific purpose and that this use of the language is actually quite sophisticated.

The third aspect of context that needs to be considered is the general educational context (Cline, 1993). Are there other LEP students in the school? How are these students incorporated into school policies, particularly testing policies? Is there an atmosphere of understanding and inclusion? Are there adults in the school who speak the same first language as the student? Are there special resources to help the student's educational progress? Does the school make an effort to reach out to the parents and guardians of LEP students? The answers to all of the questions will indicate whether the student is in an educational setting that supports his or her academic achievement. A student who is supported is likely to achieve better on a test than a student who is not (Cline, 1993).

After looking at the context of the assessment, the assessment itself should be examined for bias. According to Wilen and van Maanen Sweeting (1986), most of the available literature focuses on the description of bias rather than on solutions to it. According to Fradd et al. (1994), practitioners state that all tests are biased to some extent because they all depend on prior learning and cultural experiences. Bias cannot be eliminated, but can be minimized and analyzed carefully. Standardized tests used in large scale assessments can have several types of bias (Chamberlain & Medeiros-Landurand, 1991; Hoover, Politzer & Taylor, 1987; Taylor & Lee, 1987; Wilde & Sockey, 1995):

> **Cultural bias** (Wilde & Sockey, 1995). If test items reflect only the mainstream cultural background of the test writers and the population of students who participated in the norming, the test can be biased against students from other cultural backgrounds who have different experiences (Mercer, as cited in Chamberlain & Medeiros-Landurand, 1991). In a study of Hispanic students' performance on reading tests, Garcia (1991) found that the students performed better on reading passages that had culturally familiar content. For example, a passage on piñatas was easier for most of the students than the passages on Canada, water erosion, and chimpanzees. One way to deal with cultural bias,

especially in reading passages, would be to use a range of reading topics that involve content relating to different cultural backgrounds.

Situational bias. Taylor and Payne (as cited in Taylor & Lee, 1987) state that any testing situation is a "social occasion" with its own set of rules for communication. If a student does not follow the rules, his or her behavior might be interpreted negatively and the test score may be affected. For example, as previously mentioned, in some cultures, students may have been taught that obvious questions do not need a direct answer or that directness is not an appropriate style of communication. An unspoken rule of testing in the United States may be that the preferred style of writing is to be very direct and to give an obvious and repetitive answer to the writing prompt. In such a case, the language minority or LEP student who uses a circular style of logic in which the answer to the writing prompt can be deduced but is not stated directly, may receive a low score in spite of the fact that his or her English use was acceptable.

Linguistic bias. (Navarette & Gustkee, 1996; Taylor & Lee, 1987; Wilde & Sockey, 1995). Test items that emphasize a certain style of English rather than overall communicative ability, may be biased against students who are not familiar with that particular style of English. One example of this may be students who live in a community where a non-standard dialect of English is used but who need to use standard English on an assessment of academic content (Hoover, Politzer & Taylor, 1987; Taylor & Lee, 1987). If these students are penalized for not having had the opportunity to learn standard English, then the test is biased against them.

Communicative style bias (Taylor & Lee, 1987). Standardized test writers tend to assume that all students communicate in a similar style. For example, on a writing prompt, test writers may assume that students will write a long answer in response to a short question. Written responses that are shorter than the expected norm may receive a lower grade because they lack the expected level of detail and description. However, students from other cultures may not be aware of the American norm for length and may not be accustomed to the level of detail that is required in American academic writing. If a student is penalized for following different norms and has not been made aware of communicative style expectations, then the test has communicative style bias.

Cognitive style bias (Taylor & Lee, 1987). Research indicates that certain cultures may have a preference for a particular cognitive style due to such factors as child rearing practices and ethnic background. Standardized tests tend to require

an analytical cognitive style. Ramirez and Price-Williams (as cited in Taylor & Lee, 1987) found data to suggest that Asian and European students come from cultures that value this analytical style, but students from Hispanic and African cultures tend to prefer a different style. In addition, men and women within a cultural group may tend to prefer different styles and people from middle-class backgrounds may have different preferences from people from lower class backgrounds. Students who are not strong in the type of cognitive thinking that the test requires may not be able to demonstrate their knowledge adequately, even though they may know the answer to a test question.

Socioeconomic bias. Test items may reflect the experiences of the particular social class of the test writers (usually middle class). Politzer and Taylor (1987) give an example of socioeconomic bias that could apply equally to many African-American as well as LEP students. They cite an example of a comprehension question for a reading selection in which students were asked to supply a word to match a definition. The definition given was "If a person does something against the law, he is an:_____." The middle class test writers may have assumed that the student would pick the word "offender" from a list of possible choices. However, according to the researchers, the lower class student may have good reason to choose either the word "officer" or "official," based on the student's particular experiences. Students from other countries may also have good reason to choose another word. Many language minority and LEP students have experienced the corruption of police officers, politicians, and government officials in their native countries, so these "incorrect" answers are logical choices when the students' perspective is considered.

Test interpretation bias (Taylor & Lee, 1987). The interpretation of an LEP student's responses on a test may be biased if the responses are compared to those of a norming sample that did not include many LEP students (August et al., 1994; Chamberlain & Medeiros-Landurand, 1991; Navarette & Gustkee, 1996; Zehler et al., 1994). Taylor and Lee (1987) question the validity of norming samples because they believe that a universal norm does not exist. According to these researchers, children learn skills at a different rate and in a different order, therefore a test that expects a child to have certain knowledge at a particular age may be biased against children who do not fit the norm.

Types of Tests That May Reduce Bias

Many advocates for language minority and LEP students recommend the use of alternative assessments (e.g., observations, portfolios, and performance assessments) with these students instead of standardized tests. Estrin (as cited in Cuevas, 1996) states that alternative assessments are preferable because they are more sensitive to the needs of LEP students. Students who are being instructed in their native language can demonstrate performance in that language or in a combination of the first and second languages (Zehler et al., 1994), the tests can be adapted to the students' linguistic and cultural backgrounds (Estrin, as cited in Cuevas, 1996), and the tests should be less biased than traditional standardized tests because they allow students to respond in more than one way (Navarette & Gustkee, 1996).

In spite of the advantages of alternative assessments, educators and policymakers who consider using these assessments because of the potential for reduced bias also need to be aware that they may create several problems for language minority and LEP students:

- They may require more use of English than the standardized test; this is especially true for alternative assessments in math (August et al., 1994; Zehler et al., 1994).
- They may be more cognitively demanding, and this requires that students receive special training on how to complete such tasks (Navarette, as cited in Navarette & Gustkee, 1996; Perrone, as cited in Navarette & Gustkee, 1996; Zehler et al., 1994).
- They may not solve the problem of high exemption rates from tests; LEP students may continue to be exempted at the same rate or at an even greater rate than they currently are (August et al., 1994).
- They may have an unfamiliar format for students from other cultural backgrounds; therefore, they require that students be trained in how to take a performance assessment (Navarette, as cited in Navarette & Gustkee, 1996; Perrone, as cited in Navarette & Gustkee, 1996).

Current Large-Scale Assessment Practices with LEP Students

National Data Collection Programs

Most national data collection programs have excluded students with limited English proficiency. However, two programs provide relevant information: the National Education Longitudinal Survey (NELS:88) and the National Assessment of Educational Progress (NAEP).

The National Education Longitudinal Survey (NELS:88). The National Education Longitudinal Study of 1988 (NELS:88) was a study sponsored by the National Center for Education Statistics (NCES). It followed students who were in 8th grade in 1988 through high school to post-secondary education or work. According to Ingels (1993) and NCES (1995), NELS:88 focused on several policy issues: (1) school, classroom, family, and community characteristics associated with achievement; (2) the transition of different types of students, including LEP students, from 8th grade to secondary education, and from secondary education to either post-secondary education or to work; (3) the influence of ability grouping and "differential course-taking opportunities" on a student's educational experiences and achievement; (4) factors related to and consequences of dropping out; (5) changes in educational practice over time; (6) the schools' role in helping disadvantaged students; and (7) the academic performance and school experiences of language-minority and LEP students. Information is collected through sources such as standardized tests, parent and student surveys, attendance records, student transcripts, and teacher evaluations of student performance.

The NELS:88 study has been criticized for excluding LEP students who did not have high levels of English proficiency, and as a result, are more likely to drop out of school (Ingels, 1993; National Research Council, 1997). The exclusion of these students has the potential to skew survey results in favor of the more proficient students and to bias policy decisions that are based on the results (Ingels, 1993). However, there are still things to learn from NELS:88 regarding the assessment of LEP students. The results from this study are still being analyzed.

The National Assessment of Educational Progress (NAEP). The National Assessment of Educational Progress (NAEP) is a congressionally mandated achievement test that collects information on the performance of populations and subpopulations of students. Because it is the only assessment that samples both a national and regional cross-section of students in kindergarten through 12th-grade, information about the inclusion and participation of LEP students on the NAEP is a relevant source of information on statewide achievement of LEP students (August & Hakuta, 1994).

Inclusion/Exemption. Over the years, NAEP inclusion/exclusion criteria have changed. Before 1990, LEP students were allowed to be excluded if educators judged them incapable of "meaningful participation" in the test (NCES, 1996a, 1996b; Spencer, 1991). In 1990, the exclusion criteria became more specific. LEP students could only be exempted if they met all three of the following criteria: (1) they were non-native speakers of English, (2) they had been in an English speaking school, in classes conducted in English, for fewer than two years (time spent in bilingual classes did not count), and (3) educators judged them to be unable to participate meaningfully (National Academy of Education, 1996; NCES, 1996a, 1996b; Spencer, 1991). In a study of the 1994 testing cycle, researchers found that about 75% of the total number of students exempted could have participated in the assessment and contributed to the overall state results (National Academy of Education, 1996). More than 50% of the LEP students in the study sample were exempted even though three-fourths of these students did not meet the exemption criteria (National Academy of Education, 1996; Olson & Goldstein, 1996). A review of the inclusion/exemption policies suggested that the following factors contributed to the participation decisions for LEP students:

- NAEP officials had originally decided to exclude LEP students from the test in order to standardize the testing procedures (Olson & Goldstein, 1996).
- The policy listed criteria for identifying and excluding LEP students, rather than policies for including them (Olson & Goldstein, 1996).
- School staff believed LEP students were not able to participate meaningfully (Olson & Goldstein, 1996).
- There were few accommodations or adaptations available for LEP students who took the test (Olson & Goldstein, 1996).
- The amount of time an LEP student spent in a language program and the student's exemption from other large scale assessments negatively influenced his or her participation on the NAEP (National Academy of Education, 1996). Researchers found a pattern of excluding LEP students from statewide assessments. This pattern originated in the process of identifying and assessing LEP students for placement in a language assistance program. Students not identified as LEP did not receive services designed to meet their needs and were more likely to have low levels of achievement as a result (O'Malley & Valdez Pierce, 1994).
- School personnel interpreted the 1994 NAEP exclusion guidelines differently (National Academy of Education, 1996).

- The definition of LEP was applied inconsistently across states (Spencer, 1991).

Excluding LEP students in large numbers can have potentially serious effects. It may affect state rankings on the NAEP (McGrew, Thurlow & Spiegel, as cited in O'Malley and Valdez Pierce, 1994; National Academy of Education, 1996; Spencer, 1991), and may overestimate the effect of certain educational programs in an amount proportional to the size of the excluded population. Different interpretations of the exclusion criteria may also have positively or negatively influenced LEP performance scores (National Academy of Education, 1996).

Based on the results of the 1995 study, NAEP officials developed new *inclusion* criteria that support the inclusion of as many LEP students as possible (National Academy of Education, 1996; Olson & Goldstein, 1996; Spencer, 1991). A student must be included in the NAEP if:

- They have received the majority of their instruction in English for three or more years.
- They have received the majority of their instruction in English for less than three years but are judged to be capable of participating in the English version of the assessment with accommodations and modifications.
- They have received the majority of their instruction in English for less than three years but are judged to be capable of taking the newly developed Spanish or Spanish-English versions of the test.

When there is doubt about including an LEP student on the NAEP, educators are urged to include the student.

Accommodations. When teachers were asked about the use of accommodations for LEP students taking the NAEP, researchers found that these educators were "quite liberal" in recommending accommodations and adaptations for these students (National Academy of Education, 1996). The most commonly suggested accommodations were:

- Allowing extended time (82%).
- Creating shorter versions of tests (82%).
- Using pictures in presenting the assessment (75%).
- Providing instructions in the students' native languages (75%).
- Allowing out-of-grade testing (75%).
- Allowing individual or small group testing (52%).

If these teacher recommendations had been followed, only about 50% of the LEP students in the NAEP study sample would have been assessed under standard conditions in the 1994 Trial State Assessment (National Academy of Education, 1996).

NAEP officials wanted to keep the test as standardized as possible while providing more accessibility for LEP students. A 1995 CRESST/UCLA study on the impact of simplifying the English on standardized assessments (Abedi, Lord & Plummer, as cited in NCES, 1996b) indicated that there was no significant improvement in LEP students' math test results when English syntax and vocabulary were modified. Therefore, a translated math test in Spanish was created, because Spanish is the language of the largest percentage of LEP students in the United States. Two versions of the translated test were field tested in 1995: a Spanish-only version, and a bilingual Spanish-English version with test items in both languages (Olson & Goldstein, 1996). The results of the field test indicated that these two translated versions allowed a greater number of LEP students to participate in the assessments (NCES, 1996a, 1996b). Based on these results, translated versions were used for the first time in 1996 despite concerns about whether test results from these versions would be comparable to results from the original English version. An additional suggestion by the CRESST researchers — to examine the effects of modified semantics — has not yet been done.

State Testing Programs

Surveys of state testing policies in the U.S. often produce greatly varying results depending on the sample size, the date of the study, and from whom the data were obtained. Data from five surveys and studies are available; four studies deal specifically with LEP students. The first study was conducted by Lam and Gordon (1992). The researchers surveyed SEAs in the 50 states and Washington D.C. in 1988-89 and asked about statewide testing practices. The second study was conducted by O'Malley and Valdez Pierce (1994) in 1991. The researchers surveyed state coordinators of bilingual/ESL programs in the eastern half of the United States (including Washington D.C., the U.S. Virgin Islands, and Puerto Rico) and asked about statewide/territory-wide assessment practices. The third study is the State Student Assessment Programs survey, conducted by the North Central Regional Educational Laboratory (NCREL) and the Council of Chief State School Officers (CCSSO) (NCREL, 1996). In this section we refer to both the State Student Assessment Programs survey (NCREL, 1996) and the summary document that discusses the survey results (NCREL, 1996a, b). The fourth study (Thurlow, Liu, Erickson, Spicuzza & El Sawaf, 1996) analyzed the written testing guidelines for LEP students in assessment documents from 18 states with graduation tests. General trends from O'Malley and Valdez Pierce, Thurlow et al., and NCREL (1996a, b) are summarized here because they are the most recent and comprehensive studies of testing policies specifically for LEP students.

Participation. Educators, policymakers, and researchers disagree over whether it is appropriate and ethical to exempt students from assessments, but many states do allow exemptions (Zlatos, 1994). O'Malley and Valdez Pierce (1994) and NCREL (1996a, b) differ in the number of states they report as requiring participation of LEP students in state assessments. In a 1991 study, O'Malley and Valdez Pierce reported that 31 of 34 states and territories in the eastern U.S. required all students to participate in statewide assessment. In contrast, NCREL (1996a, b) reports that 36 of 50 states allowed LEP students to be exempted from such assessments. While the sample sizes for the surveys differ, it is clear that there is a sizable difference in the number of states reporting the required participation of LEP students versus those that allowed LEP students to be exempted. In O'Malley and Valdez Pierce's 1991 survey of the 31 states with statewide testing programs, six states with graduation testing had an alternative plan for LEP students who did not pass the test. Four out of these six gave failing students a certificate of attendance, and 2 of 31 provided assessments in the students' native language if the student did not pass the English version.

Assessment Guidelines. Practices in testing LEP students are affected by the availability of clear and rational assessment guidelines (Lam, 1993; Lam & Gordon, 1992). In general, states with well developed guidelines at least have given thought to their policies. In the area of test guidelines for LEP students, O'Malley and Valdez Pierce (1994) found that 34 states had state guidelines or requirements for testing LEP students, but that these guidelines were so broad and there was so much flexibility in their interpretation that they were of little value to staff in making assessibility decisions. The NCREL study (1996a, b) also found a great deal of local flexibility in the interpretation of guidelines. To determine assessibility, 6 of 34 states used English proficiency level, but survey respondents from these six states did not give discrete cut off scores on specific proficiency tests. Three states (3 of 34) tested LEP students but excluded the scores from statewide reporting. NCREL (1996a, b) reported that the majority of states (n=50) used either the number of years in an LEP language program or the second language proficiency as the criteria for determining assessibility. However, no specific numbers were given for these data. Even though the data from the two studies are not easily comparable, it appears that a majority of the states have some type of LEP assessment policy. Lam and Gordon (1992) report that there is a general trend for states receiving bilingual education funds to have LEP student testing policies and that the percentage of LEP students in a state may also positively correlate with the presence or absence of LEP student testing policies. The researchers suggest that the more LEP students a state has, the more likely that state is to have a testing policy for them.

Modifications and Accommodations. O'Malley and Valdez Pierce (1994) report that most states in their study did not recommend the use of test modifications. The only accommodation that they found being used more than once was testing LEP students separately with a familiar examiner (2 of 34 states). In contrast, NCREL (1996a, b) found that while seven of 50 states include LEP students in assessments with no accommodations, 25 of 50 states reported that they did allow accommodations for LEP students. Of those 25 states, only 17 responded with the specific accommodations allowed. They identified:

- Separate setting (17 of 17).
- Flexible testing schedule (15 of 17).
- Small group administration (15 of 17).
- Extra time (14 of 17).
- Simplified directions (11 of 17).
- "Other" unspecified accommodations (10 of 17).

Some of the less popular accommodations were the use of other languages on the test (4 of 17) and the use of an alternative test for LEP students (3 of 17).

In the area of programs and services to help LEP students pass state assessments, O'Malley and Valdez Pierce (1994) reported that 3 of 34 states in their survey emphasized the teaching of test taking strategies so that students would not require accommodations. Thirteen states (13 of 34) had guidelines to provide testing information to parents in English. Only 8 of 34 states had guidelines requiring this testing information to be translated into the parent's native language.

High Stakes Graduation Testing. As of 1994, 18 states had high stakes graduation testing programs in place (Bond & King, 1995; Thurlow, et. al, 1996). In these states, students are required to obtain a minimum score on the test to obtain a diploma. The majority of the states (17 of 18) used criterion referenced tests with cut off scores determined by SEAs (Klein, as cited in Bond and King, 1995). One state used both a criterion and a norm-referenced test and one state was switching from a norm referenced to a criterion referenced test. All programs relied primarily on multiple choice test items due to the high level of technical quality required in high stakes testing (Bond & King, 1995). The average testing program assessed 10th or 11th grade students on math (17 of 18 states), reading (14 of 18 states), and writing (13 of 18 states) to determine whether students had the basic skills needed to graduate (Bond & King, 1995).

The Existence of Written Testing Guidelines. Sixteen of the 18 states had some form of written testing guidelines for LEP students (Thurlow et al., 1996). Most of the guidelines were

written for students in special education programs and *LEP students combined; it was often* unclear which guidelines applied to LEP students and which ones did not. Most of the guidelines were vague and general, providing little information on how to make testing decisions for LEP students.

According to Thurlow et al. (1996) three states had documents describing how inclusion/ exemption decisions were made for LEP students. These states used a combination of a language proficiency score and the length of time in the school system to determine assessibility. North Carolina allowed LEP students to receive an exemption from only one subject test. Maryland gave its LEP students a native language proficiency test first and if a student obtained a score below the cutoff, the student could be exempted from one administration of each segment of the test. In New Jersey, LEP students who enter the school system in grade 9 or later can be exempted from the graduation test. Exempted LEP students are required to take a native language proficiency test and if their proficiency is above the minimum level they must then take the Maculaitis Assessment program and get a specified score in order to be eligible for a state endorsed diploma. Two states allowed LEP students to use an alternative procedure in order to demonstrate their skills. Four states have more than one type of diploma; often one is based on the number of credits a student receives and the other is based on a state exam.

Accommodations on Graduation Tests. Nine out of 18 states did not list any testing accommodations or adaptations that were allowed for LEP students on their graduation tests (Thurlow et al., 1996). The remaining nine states used the following accommodations:

- Extra time (5 of 18).
- Separate room (5 of 18).
- Bilingual dictionary/electronic translator (4 of 18).
- Oral translation of test (1 of 18).
- Written translation of test (1 of 18).
- Written translation of directions (1 of 18).
- Marks in test book (1 of 18).

Strategies for Supporting Students Who Do Not Pass High Stakes Assessments. For students who do not pass high stakes graduation exams, there are two types of strategies that schools use to improve performance (Bond & King, 1995). First, many schools work on increasing student and parent awareness of the test. Catterall (as cited in Bond & King, 1995) studied several

states with graduation tests and found that only about 50% of students in those states were aware of the tests despite the serious impact the tests could have on their future. To increase awareness, some schools have experimented with cable television programs on test preparation while others have scheduled the test on Saturdays so that the entire community is aware of the testing program and students have no other distractions.

The second type of strategy that schools use to improve test performance is remediation (Bond & King, 1995). There are currently two types of remedial programs: tutoring programs and self-tutoring programs. Some schools have developed a one-to-one tutoring program using university students as the tutors. Other schools have enhanced their computer laboratories and created computerized tutorial programs specific to the content areas tested so that students can work individually on the areas that need improvement. The key to developing high quality remedial programs for students is to have adequate funding. For example, 7 out of 18 states have special funds for remedial programs: two states use Title I or other compensatory funds, one uses unspent GED/Adult education funds, and one includes money for remediation in its funding formula (Bond & King, 1995).

Recommendations for Graduation Testing Programs. Based on their review of written testing guidelines for high stakes assessments, Thurlow et al. (1996) make the following recommendations for assessment guidelines:

- Put guidelines for high stakes graduation exams in a separate section from guidelines for other types of tests.
- Define the term "LEP" with specific criteria so districts/schools interpret the term in the same manner.
- List guidelines for LEP students separately from those for special education students.
- Explain how assessibility is determined and give clear criteria.
- Explain who should be involved in making assessibility decisions.
- Clearly define each accommodation and give examples of it.
- Say how scores will be reported and indicate whether scores for LEP students will be disaggregated.

What the Literature Recommends

Several recommendations about large-scale assessments and students with limited English proficiency can be gleaned from the literature. Some of these recommendations are summarized in this section. The list should not be considered exhaustive, nor should the reader assume that we necessarily endorse them.

Test Development

- Create a universal standardized definition of the term LEP. This would allow results to be compared directly because they would be from the same student populations (NCES, 1996b; Spencer, 1991).
- Consider developing levels of tests related to a student's English proficiency. Each level would have different tasks (National Academy of Education, 1996).
- Design the test/data collection effort to include as many LEP students as possible (Ingels, 1993; NCES, 1996b). Make sure that the assessment includes items at lower difficulty levels so that the test can measure more accurately the growth in performance of low-achieving LEP students over time (NCES, 1996b).
- Develop "normative conceptions" of progress in educational achievement for students receiving bilingual services and LEP services (NCES, 1996b, p. 37).
- Consider using item formats other than multiple choice. For example, turning a multiple choice question into an essay question allows more students to demonstrate their knowledge (Ingels, 1993; NCES, 1996b).
- Consider using nonverbal tests to collect achievement data in some subjects. Tests requiring a large English vocabulary may not give an accurate picture of achievement because LEP students' knowledge may exceed their English ability (Ingels, 1993; NCES, 1996b).

Test Administration

There is now general agreement that students with limited English proficiency should participate in large-scale assessments as often as possible. Yet, there is a practice of delaying tests for LEP students in the belief that more time will allow these students to learn more English and to achieve at a higher level. Zlatos (1995) discusses the different forms these delays can take:

- The student is failed and repeats the same grade so that he or she takes more time to reach the grade where testing occurs.
- The student is put in a "transition" grade where he or she takes more than one year to finish one grade.
- The student is assigned to a remedial class where students are not required to be assessed.

These practices often are followed in the belief that the students will be helped, but the end result may be lowered expectations for the students. The literature suggests that all of these practices can have significant negative consequences. For example, there are many potentially negative outcomes of failing a student in order to delay testing. Slavin (as cited in Zlatos, 1995 p. 101) states that flunking "inflates subsequent retesting scores by a significant percentage." Le Mahieu and Haycock (as cited in Zlatos, 1995) caution that children who are retained a grade may receive teaching unsuited to their needs and these students may be behind their peers for many years as a result. Students who are retained may have a higher tendency to drop out of school.

Other potentially negative results that occur when students are excluded from assessments are:

- Students do not get practice taking the tests and do not develop familiarity with them (Rivera & Vincent, 1996).
- Students needing remedial services are not identified until it is too late to help them (Rivera & Vincent, 1996).
- Districts and schools remain ignorant of the needs of LEP students (Lacelle-Peterson & Rivera, 1994; Sosa, as cited in Zlatos, 1994).
- LEP students who have no alternative ways to demonstrate their knowledge may not receive the benefit of state and federally funded language services and other programs (O'Malley & Valdez Pierce, 1994). This may happen even if the students are already identified as limited English proficient.

Among the positive effects that occur when LEP students are included are:

- The overall results are more comprehensive (Zehler et al., 1994).
- The comparisons made between schools, districts, and states are more appropriate and fairer if the same population of students is tested in each group (Zehler et al., 1994).

- There is more pressure on schools and districts to improve LEP services if student test scores are made public (Zehler et al., 1994).

- There are more accurate data available for further study. On other types of assessments there is general consensus in the field to assess as many LEP students as can practically be included so that schools can be held accountable for the students' educational progress (August et al., 1994; Haycock, as cited in Zlatos, 1994; Resnick, as cited in Zlatos, 1994; Zehler et al., 1994).

Determining Assessibility

Having said that LEP students should participate as often as possible, one major difficulty remains: how should LEP students' assessibility be determined? There is disagreement over the criteria for inclusion and exemption. For example, according to Lam (1993), in 1980 states usually based exemption decisions on language proficiency or parent and teacher recommendations. However, Rivera and Vincent (1996) say that most states now use proficiency or time in the system to determine assessibility for tests other than graduation tests.

In general, four schools of thought address the topic of assessibility:

1. **Assessibility should be based on the number of years the LEP student has been in the American educational system.** These educators generally recommend that a student be in English speaking American schools for 1-1/2 years (time enough for some English instruction) before participating in large scale assessments (August et al., 1995; Zehler et al., 1994). One problem with basing participation on time in the school system is that there are no allowances made for students' different rates of English acquisition (Gandara & Merino, 1993). Depending on their backgrounds and individual characteristics, some students may acquire very little English in 1-1/2 years and some may acquire a great deal. A related issue is that 1-1/2 years is not a significant amount of time when one considers how long it takes a language learner to acquire academic competence in a second language (Rivera & Vincent, 1996). The amount of time that LEP students have to study in high school and obtain their diploma may not be sufficient to completely develop academic language proficiency before learning academic content (Saville-Troike, 1991). Therefore, students must learn English and academic content at the same time, and learning content may require that students be tested on their knowledge.

2. **Assessibility should be based on a student's general level of English proficiency** (August et al., 1995; CCSSO, 1992a; NCES, 1996b). This criterion allows for differential rates of second language acquisition but it does not take into account the fact that proficiency tests are not always a good indicator of student performance on content area tests for several reasons:

    - A proficiency test usually includes measurements of oral/aural language and this type of proficiency is not a factor in reading and writing tests of academic content knowledge (Lam, 1993).

    - Proficiency tests measure English proficiency in a variety of ways, depending on the underlying theory. Additionally, the cut off points for determining high and low proficiency vary from test to test and from school district to school district. Thus, a student may be unable to take a statewide assessment because of the particular proficiency test taken or because of the particular cut off score used in the district where the test was taken (Lam, 1993).

    - A proficiency test does not demonstrate a student's content knowledge. It only shows the student's ability to express that knowledge in English (Saville-Troike, 1991; Mercado & Romero, as cited in Cuevas, 1996).

3. **Assessibility should be based on student, parent, and teacher judgment of the student's skills.** This judgment may often be biased because of the parents' lack of English proficiency, and the lack of reliability in information that students self-report (NCES, 1996b).

4. **Assessibility should be based on a student's literacy level in English** (see Garcia, 1991). NAEP researchers have decided to use this criterion for determining participation because it measures a student's proficiency in reading English instead of measuring other types of proficiency that are not involved in statewide assessments (NCES, 1996b).

Inclusion/Exemption

The literature makes the following recommendations for inclusion and exemption decisions:

- Determine assessibility based on English proficiency rather than on time spent in school (NCES, 1996b; O'Malley & Valdez Pierce, 1994). Have very clear

criteria (e.g., a specific score on a specific test) that accounts for a student's ability to read and write in English since these skills are required on large scale assessments (Lam, 1993). Time in school does not predict test performance accurately (NCES, 1996b). Note: due to the amount of controversy over the nature of proficiency and how to assess it, researchers have widely varying recommendations on this point. See the following articles for specific recommendations: Bracken & Barona (1991); CCSSO (1990, 1992); Gandara & Merino (1993); Guerrero & DelVecchio (1996); Lacelle-Peterson & Rivera (1994); Lam (1993); O'Malley & Valdez Pierce (1994); and Shinn and Tindal (1988).

- Use more than one source of data to make inclusion/exemption decisions (Cuevas, 1996; Ingels, 1993; Lam, 1993). Consider using length of time in an ESL program, teacher rating of English proficiency, and transcripts (Cuevas, 1996; Ingels, 1993). Collect indirect performance data on all LEP students from the following people (CCSSO, 1992a, 1992b; Ingels, 1993; NCES, 1996): bilingual/ESL teachers, regular teachers, school and program administrators, parents, and school records.

- Collect data on excluded students and periodically reassess their eligibility to participate based on the data (August et al., 1994; Ingels, 1993).

- Administer proficiency tests regularly to those who do not achieve a high enough score the first time (CCSSO, as cited in Gandara & Merino, 1993; CCSSO 1992a, 1992b; NCES, 1996). The CCSSO recommends administering the tests annually.

- Include LEP students in assessments for accountability even when there is doubt about the student's ability to take them (Olson & Goldstein, 1996; O'Malley & Valdez Pierce, 1994; Zehler et al., 1994).

- Establish a "gray area," a range of scores on a proficiency test that indicates the student should have his or her situation reviewed by a committee. The committee would determine whether the student should participate in the assessment (Lam, 1993).

- Use an alternative method to monitor exempted students' academic progress (August, et al., 1994; August et al., 1995). Performance assessments are recommended because they allow a greater range of students to participate (Ingels, 1993; NCES, 1996b).

- Work on including a larger percentage of LEP students in the test (Spencer, 1991). Take into account that the decision to exclude these students is often based on the fact that it costs more to assess them.

- Specify how inclusion decisions should be made and create a clear decision-making tree. Leaving decisions up to schools and states results in a lack of standardization and therefore a lack of comparability (NCES, 1996b).

- Develop criteria for inclusion instead of focusing on exclusion (Olson & Goldstein, 1996).

- Set time limits for exemption. A student should not be exempted indefinitely (August et al., 1994; NCES, 1996b)

- Collect information about excluded students even if it is not comparable to test results (e.g., teacher ratings, imputed scores, etc.) (August et al., 1994; NCES, 1996b).

Accommodations and Modifications

Written Translations. One of the least commonly used but much discussed accommodations for LEP students is the use of translated test documents on statewide assessments. Translations can either be written ahead of time, or interpretations can be done orally at the time of testing. Either the entire test may be translated or interpreted, or this can be done for only one piece, such as test directions.

There is general agreement that students should be allowed to take an assessment in a language that will not interfere with their performance. This opinion might not hold true for educators. At a July 1996 focus group meeting of LEP educators conducted by the Minnesota Department of Children, Families and Learning (Spicuzza, Erickson, Thurlow, Liu & Ruhland, 1996), the educators' general belief was that LEP students should be required to take the high stakes graduation test in English. They believed that a translated or interpreted test was inappropriate because students would be required to use English in work and in higher education. If a translation or interpretation is used, careful consideration must be given to its appropriateness for a particular student (Wilen & van Maanen Sweeting, 1986).

Written translations may benefit two types of LEP students: (1) students who received grade appropriate instruction or educational experience in their first language or in a bilingual program (August et al., 1995; Rivera & Vincent, 1996; Zehler et al., 1994), and (2) students who are more fluent in their first language than their second, even though they have not been instructed

in their first language, and who choose to take a translated version (August et al., 1994). It is worth noting, however, that there are few high schools offering bilingual education services, so most LEP students will not have had access to the content of the test in their first language unless they received it in their native country (Rivera & Vincent, 1996).

It is important to involve students in the decision-making process because some students may not be literate in their first language even though they are orally proficient in it (DeAvila & Havassy, as cited in Wilen & van Maanen Sweeting, 1986). A case in point are the Hmong students in the United States. Because the Hmong writing system was only developed within the past few decades, many Hmong parents may not be literate in their native language. As a result, the students may have had little exposure to the writing system even though they communicate in Hmong with their parents and with others in their ethnic community. A written translation of a test may not benefit a student in this situation. Another type of student who may not be helped by a written translation is one who speaks a combination of two languages as his or her first language (Wilen & van Maanen Sweeting, 1986). Such a student might come from the Chicano community where a combination of Mexican Spanish and English is a form of language that many children use as their first language (Penalosa, 1980).

Negative Aspects of Written Translations. If a state or district is considering the use of a written translation of a test, there are several negative aspects that need to be taken into account. The major drawback to written translations is the technical problems associated with them (Olmedo, 1981). The first technical problem is the quality of the written translation. Test directions may be extremely difficult to translate into a second language because they are too technical (Bracken & Barona, 1991; Zehler et al., 1994). Test items may be difficult to translate as well because a concept that is relatively easy in English may be much more difficult in another language, if the concept exists at all (Fradd et al., 1994). When vocabulary items are translated into a second language, their level of difficulty may change if the words are infrequently used in the other language (Bracken & Barona, 1991; Fradd et al., 1994; Rivera & Vincent, 1996; Wilen & van Maanen Sweeting, 1986). This increased difficulty may be especially important on vocabulary tests that require knowledge of synonyms, antonyms, and analogies, or on cloze tests that require students to supply a missing word from a sentence (Rivera & Vincent, 1996). Test items may also be difficult to translate because translating the words in the item does not guarantee that the meaning of the item will be the same in the second language as it was in the first (Bracken & Barona, 1991; NCES, 1996b; Olmedo, 1981; Rivera & Vincent, 1996; Wilen & van Maanen Sweeting, 1986).

The second technical problem is that students who speak the same native language may speak different dialects (Bracken & Barona, 1991; Nutall, 1987; Olmedo, 1981); therefore, a test that is written in Mexican Spanish, for example, may be difficult for a child who knows another dialect of Spanish. Certain vocabulary items may differ from one dialect to the other and the

student may not comprehend test items as a result, or the student may answer items using his or her own dialect and these answers may be misinterpreted by the scorer (Wilen & van Maanen Sweeting, 1986).

A third technical problem is that reliable translations may not be available for dialects or languages (August et al., 1994). As an example, in California there are 96 languages spoken by LEP students, and other states may deal with similar numbers of different languages. Qualified translators may be unavailable for many of them (Fradd et al., 1994). Educators express concern about the legality of offering written translations to some linguistic groups and not to others (Spicuzza et al., 1996) and as a result they are cautious about giving a translated test to any student.

A fourth technical problem is that a translated version of an English standardized test may not be renormed on the population for which it is intended (Fradd et al., 1994; Nutall, 1987; Wilen & van Maanen Sweeting, 1986). Some translated tests still rely on the norms from the English version.

Recommendations for Written Translations. If a state or school district chooses to use a translated version of a test, the following recommendations are provided:

- Understand that translated test versions are problematic due to such factors as the difficulty of creating an accurate translation of a test, the variety of first and second languages spoken by LEP students, the amount of instruction LEP students receive in English, a lack of literacy in the native language, and a lack of instruction in the native language (NCES, 1996b).
- Make sure that the words in a translation are as common as the words on the original test and that they have the same level of difficulty (NCES, 1996b).
- Make sure the translation is high quality (Rivera & Vincent, 1996).
- Consider translating the instructions instead of the entire test. Often, the directions for a standardized test are more difficult than the test questions. If a student misunderstands the directions, he or she can do poorly on the entire assessment even if he or she knows the content (Zehler et al., 1994).
- Conduct studies to prove that the translation is equivalent to the original version (Bracken & Barona, 1991).
- Outline a decision making process and the criteria used to determine whether it is appropriate to test a student with the translated version. Decisions should also

be made about the scheduling of the translated version, how to use it, and when a student should be required to take the English version (Zehler et al., 1994).

- Cooperate with other states or school districts. Share assessments in other languages if they test appropriate content or consider having a shared item bank from which many states or districts can draw test items. This will cut down on the burden for an SEA or LEA to develop assessments in all the LEP students' languages (August et al., 1994; August et al., 1995).

- Keep in mind that even if a translation is helpful for some students who have literacy in their first language, the cultural experiences and individual differences of these students may still make the assessment unsuitable for them (Bracken & Barona, 1991; Figueroa, 1990).

- Choose words for the English version of the test that will translate clearly into another language (NCES, 1996b). Make modifications to the English version to accommodate the translated version. Werner and Campbell (as cited in Bracken & Barona, 1991) recommend writing simple English sentences, repeating English nouns instead of using pronouns (pronouns often do not translate clearly into a second language), avoiding metaphors and colloquial English, and avoiding "hypothetical phrasing and subjunctive mood" in English (p. 120).

Oral Interpretations. There is not much information in the literature on the topic of oral translations of tests. NCES (1996b) found that in a study of the CLAS test in California, some interpreters unintentionally coached students on the test items while translating the directions. For this reason, educators are cautioned against using on-the-spot interpreters in psycho-educational assessments (Nutall, 1987). It is recommended that there be a standardized script in the second language that the interpreter reads from and explains to the student (Figueroa, 1990; Wilen & van Maanen Sweeting, 1986). Psycho-educational tests frequently include an oral language assessment, while large scale standardized tests do not. However, the finding of unintentional coaching on an oral interpretation may have implications for large scale testing.

Other Types of Accommodations and Modifications. In addition to translated and interpreted versions of tests, there are other types of modifications and accommodations that can be made for LEP students. The major advantage in using other testing options for these students is that they encourage the students to take the test in English (August et al., 1994; Ingels, 1993). In a study of the NAEP, the National Center for Education Statistics (1996b) found that when there were more accommodations available for LEP students, teachers encouraged more of the students to participate in the test. On other types of tests, such as state tests, some educators believe that accommodations do not increase the participation or accuracy of results for many LEP students

(Rivera & Vincent, 1996). Some students with higher levels of English proficiency may be helped by accommodations, but students with little English proficiency need more help than accommodations can give (Rivera & Vincent, 1996). For example, some accommodations are based on the assumption that a student is literate in his or her native language. The researchers state that allowing the use of a bilingual English-native language dictionary presupposes that the student is literate in the first language and in English. This implies that the student only requires a translation of the content word into English because he or she knows the content words in the native language. Such an assumption may not be true for students who have not studied the academic content in their native language or do not have sufficient literacy skills to use the bilingual dictionary.

Research Needs on Accommodations. There is little research that shows how accommodations for LEP students affect test validity. The literature calls for research to be done on the following types of accommodations:

- Changing the test content so that it is directly relevant to the background of the student (August et al., 1994).
- Reading the test to the student in English or having an audiotaped reading (August et al., 1995; Ingels, 1993; NCES, 1996b). These options may apply to the entire test or only to the directions.
- Offering student response options such as drawing pictures, audiotaping answers, using the native language to respond, dictating answers to a scribe, etc. (August et al., 1995; NCES, 1996b; Wilde & Sockey, 1995).
- Providing clarifying information at the end of the test booklet (e.g., definitions of words, etc.) (August et al., 1995; NCES, 1996). Does it provide students with an unfair advantage because it assists them with content? See the section on NAEP for a description of the research on this accommodation.
- Simplifying the English on the test (e.g., paraphrasing test items) (August et al., 1994, 1995; Ingels, 1993).
- Using computer assisted assessments adapted to meet the needs of LEP students (August & Hakuta, 1994; August et al., 1994) and allowing on-line synonyms for difficult vocabulary (NCES, 1996b).
- Pretesting students before the assessment is given. Pretesting would identify students with learning problems and give the students experience taking the test (Walstad, 1984). Research has shown that students may perform poorly on

assessments if they have a lack of familiarity with the test (Lam, 1993). According to Walstad (1984), school districts that pretested their students showed a large increase in scores.

- Using an examiner familiar to the students (Fuchs & Fuchs, 1989). The research done on this accommodation looked at the effect of examiner familiarity on minority students, but it did not control for the variable of socioeconomic status. Therefore, research cannot conclusively say that a familiar examiner improves the test performance of all LEP students, but it is believed to be helpful.

- Allowing the use of aids such as dictionaries and calculators (Ingels, 1993).

- Allowing extra time or modified testing schedules over the course of several days (Ingels, 1993; NCES, 1996b). While extra time is commonly used, the effects of giving it are not clearly understood (Rivera & Vincent, 1996). In a study of SAT and GRE scores for students who received extra time on the tests (see NCES, 1996b for a description), the extra time appeared to bias the results and overpredict students' grades in postsecondary education.

- Allowing small group or individual test administration (Ingels, 1993).

- Allowing the test administrator to use visuals when presenting the instructions to the students (Ingels, 1993).

- Allowing the use of "think alouds" (i.e., orally expressing thought processes) (August & Hakuta, 1994).

- Allowing students to dictate answers to a scribe (Ingels, 1993).

Recommendations for Using Accommodations and Modifications. Some researchers recommend allowing accommodations and modifications that can increase participation rates (Ingels, 1993; NCES, 1996). If an SEA or an LEA decides to offer accommodations and modifications on a large scale assessment, the literature has several recommendations:

- Avoid using accommodations for students with disabilities as the standard of comparison for the types of accommodations offered to LEP students. Accommodations used for students in special education are not necessarily helpful to students whose limited English proficiency prevents them from having access to the content of the assessment (Rivera & Vincent, 1996).

- Consider what types of accommodations the student receives in the mainstream classroom and use the same ones for the testing situation (Ingels, 1993). Keep in mind that any accommodations for LEP students should focus on reducing the language load of the test (August & Hakuta, 1994).

- Create modifications that are not a burden to teachers (NCES, 1996b).

- Develop a range of allowable accommodations for students of differing proficiency levels. Allowing only one type of accommodation or modification will not help all LEP students (August et al., 1995; August & Hakuta, 1994).

- Give the unmodified version of the test only to LEP students who have enough English proficiency (NCES 1996b).

- Make test instructions explicit (NCES, 1996b).

- Examine the language in the directions and test items. Modify the language when it is appropriate to do so (NCES, 1996b). Consider developing different test versions with the same content written at different English levels. A student with lower proficiency could take a version with simplified English (National Academy of Education, 1996).

- Use "less conventional" methods to test students who cannot take an unmodified version (e.g., some kind of adaptation of the test) (NCES, 1996b).

- Evaluate accommodations to determine how many students they actually help and to what degree the students are helped (Rivera & Vincent, 1996).

- Remember that accommodations and modifications do not eliminate all of the barriers that LEP students experience on tests. There may be students who are not helped at all by the types of accommodations and modifications that are allowed (Rivera & Vincent, 1996).

Scoring, Data Use, and Reporting

For the purpose of clarity, recommendations in this category are divided into four stages: before the test, after the test — short term, after the test — long term, and presenting data to the public.

Before the Test:

- Address the issue of how to score translated test versions; this includes the use

of code switching (i.e., the use of two languages in the same response) (NCES, 1996b).

- Determine whether all students have access to the content of the test (Rivera & Vincent, 1996).

- Consider characteristics of LEP students and of second language acquisition when developing scoring rubrics and rating scales (NCES, 1996b; Rivera & Vincent, 1996). Rubrics may need to be translated into the second language (NCES, 1996b).

- Do a cost-benefit analysis of different methods for determining error rates. The method chosen can influence the size of the student population that falsely receives a passing score (Huynh, 1990).

After the Test — Short Term:

- Consider imputing scores for students who are exempted from large scale assessments (NCES, 1996b). Imputed scores can be based on grades or performance on alternative assessments. Keep in mind, however, that imputed scores may skew the test results if they are not calculated accurately.

- Give each untested student a score of 0 instead of imputing a possible score. Add the 0 in with the other scores (August et al., 1994).

- Consider adjusting the scores for students who have been kept in the same grade for several years. These students often achieve higher because they have studied the same material for another year; the results are not an accurate indicator of school performance or of student progress (Slavin, as cited in Zlatos, 1995).

- Impute missing test data in cases where students do not complete all sections of the test. This is especially valuable if auxiliary data have been collected (NCES, 1996b).

- Disaggregate the data by LEP status (NCES, 1996b) and by the content area (August et al., 1994, 1995). Do this at the school, district, and state level. If individual students are identifiable in school-level data, then report disaggregation at the district level (August et al., 1994; August et al., 1995; Ysseldyke, as cited in Zlatos, 1995; Zehler et. al., 1994). This type of disaggregation allows for description of LEP student performance at the group level and inferences can be made about whether students are making progress toward standards. Consider further disaggregation by the type of language support services received or the

level of second language proficiency (August et al., 1995). It should be noted that not all researchers agree that disaggregating data by LEP status is beneficial. August and Hakuta (1994) state that doing so perpetuates the lower class status of LEP students by making them different from the other students. Other researchers believe that data should be disaggregated only for those LEP students who are tested under non-standard conditions (National Academy of Education, 1996).

- Disaggregate data by the type of accommodation that LEP students received (NCES, 1996b).

- Disaggregate by former LEP status so that these students can be tracked (August et al., 1995; CCSSO, 1990; NCES, 1996b; Rivera & Vincent, 1996).

- Disaggregate by socioeconomic status so that results related to this variable are not confused with results related to LEP status (August et al., 1995; Rivera & Vincent, 1996).

<u>After the Test — Long Term:</u>

- Document the performance of LEP students who take the test several times, including students who eventually drop out (Rivera & Vincent, 1996).

- Determine the percentage of overall change in test scores for LEP students from one year to the next rather than relying only on the absolute level of performance within a year (August & Hakuta, 1994).

- Monitor the longitudinal test data on LEP and former LEP students in order to determine whether system-wide reforms or remediation are having an effect (O'Malley & Valdez Pierce, 1994).

- Analyze the performance of exempted students and compare it to the performance of included students (Spencer, 1991).

<u>Presenting Data to the Public:</u>

- Remember that the impact of the test depends largely on the way in which the results are used (O'Malley & Valdez Pierce, 1994).

- Report scores as either pass or fail if the test was not designed to be diagnostic (Mehrens, 1993). Give failing students some information about how close they were to passing and what general areas they need to work on. This method is

generally preferable to telling students how many items they had correct (the raw score).

- Make sure that when disaggregated results are made public, factors affecting LEP students' performance are discussed (NCES, 1996b). Publishing only the results without the necessary background information could give the public the wrong idea about the capabilities of the students and schools (NCES, 1996b). Inform the public of the following points:

  1. Scores for LEP students are an outcome of English proficiency AND content knowledge. Academic achievement is being measured in the students' second language and does not necessarily indicate what they have learned in the first language (Zehler et. al., 1994).

  2. Scores may reflect the students' lack of opportunity to learn more than they reflect English ability (NCES, 1996b).

Test Preparation

Spicuzza et al. (1996) conducted focus groups with teachers of LEP students and with school administrators in 1996. They found that many teachers wanted LEP students to receive more training in test taking skills before taking the statewide assessment required for graduation. The search of the literature did not turn up many resources for this, but the work of Lam (1993) does make reference to both test taking skills and motivational strategies.

Test Taking Strategies. Offer test training classes for standardized tests. A lack of knowledge about how to take a test can reduce the validity of the test scores for LEP students (Lam, 1993).

Motivational Strategies. Use motivational strategies to improve students' motivation and to reduce stress (Haladyna et al., as cited in Lam, 1993):

- Provide extra snacks and meals.
- Give rewards for completing sections of the test.
- Send notes home to parents on how they can help their student prepare for the test.
- Talk to students about the importance of performing well on the test.

- Make changes in the school routine to show students how important the test is.
- Encourage good attendance during testing week.

Summary

In conclusion, this literature review has highlighted the following issues in the assessment of linguistic minority and LEP students:

1. Large-scale assessments for achievement or accountability are an established part of American education and the trend is not likely to disappear in the near future. The majority of states have these types of tests and of those that do not, most have indicated that they are developing tests.

2. Testing has an important role in our society and serves many purposes. In recent years, there has been more emphasis on the appropriate assessment of students who have not been well served by the educational system in the past. Language minority and LEP students are only one such group, but they are a key group. Appropriate assessment of these students means they should be included in assessments to the greatest degree possible.

3. The language minority and LEP student population is growing rapidly and the needs of these students must be directly addressed in the design and implementation of large-scale assessments. The students have unique needs that go beyond language learning and if they are to succeed, educators must begin to address those needs now.

4. One crucial factor in the debate over assessing LEP students is the lack of a standardized definition for "limited English proficient." States and school districts do not consistently apply the same term to the same population, this in turn creates difficulty in addressing the needs of the students.

5. Schools and states should develop clear guidelines for the testing of LEP students.

6. Assessibility of language minority and LEP students should be based on their English reading and writing proficiency since the tests involve reading and writing.

7. Schools and states should have a system for keeping track of the educational progress of exempted students.

8. When language minority and LEP students are tested, certain factors that affect test scores must be taken into account. The degree of acculturation, the level of first and second language proficiency, the extent of cognitive development and literacy in the native language, attitudinal factors, and test bias are just some of these factors.

9. Currently, there is little information on the effect of allowing language minority and LEP students to use specific testing accommodations. There is a great need for this type of research to be done. Research could help to determine which accommodations are most beneficial to students yet do not change the rigor of the testing situation.

10. Research from NAEP and NELS:88 addresses issues relating to the assessment of LEP students. It is a valuable reference for states and school districts.

11. States need to pool their knowledge about LEP students and their assessment resources. The demands of a testing program that focuses on the inclusion of LEP students can be great, and states may be better able to deal with these demands if they work together.

References

Adamson, H.D. (1993). *Academic competence—Theory and classroom practice: Preparing ESL students for content courses*. NY: Longman Publishing Group, 10-42.

Airasian, P.W. (1987). State mandated testing and educational reform: Context and consequences. *American Journal of Education*, *95*, 393-413.

Anstrom, K. (1996, summer). Defining the LEP student population. *Directions in Language and Education* [On-line]. Available: http://www.ncbe.gwu.edu/ncbepubs/directions

Anstrom, K. & Kindler, A. (1996). *Federal policy, legislation and education reform: The promise and the challenge for language minority students* [On-line]. Available National Clearinghouse of Bilingual Education Resource Collection Series No. 5.: http://www.ncbe.gwu.edu/ncbepubs/resource/fedpol.html

August, D. & Hakuta, K. (1994). Evaluating the inclusion of LEP students in systemic reform. In *Issues and strategies in evaluating systemic reform.* Papers prepared for the U.S. Department of Education, Office of the Under Secretary, Planning and Evaluation Service.

August, D., Hakuta, K., Olguin, F., & Pompa, D. (1995). *LEP students and Title I: A guidebook for educators* [On-line]. Available NCBE Resource Collection Series: http://www.ncbe.gwu.edu/ncbepubs/resource/lepguide

August D., Hakuta, K., & Pompa, D. (1994). *For all students: LEP students and Goals 2000. A discussion paper.* Paper presented at the NAE panel meeting No. 15., Washington, D.C.

Bond, L.A. & King, D. (1995, Nov.) *State high school graduation testing: Status and recommendations.* Oak Brook, IL: NCREL.

Bracken, B.A. & Barona, A. (1991). State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *School Psychology International*, *12*, 119-132.

Chamberlain, P. & Medeiros-Landurand. (1991). Practical considerations for the assessment of LEP students with special needs. In E.V. Hamayan and J.S. Damico, (Eds.) *Limiting bias in the assessment of bilingual students* (pp. 122-156). Austin, TX: Pro-Ed.

Chamot, A.U. (1992, August). *Changing instruction for language minority students to achieve national goals* [On-line]. Paper from the Third Plenary Session of the Third National Research Symposium on LEP Student Issues, Washington, D.C. Available: http://www.ncbe.gwu.edu/ncbepubs/symposia/chamot.html

Cline, T. (1993). Educational assessment of bilingual pupils: Getting the context right. *Educational and Child Psychology, 10* (4), 59-68.

Cloud, N. (1991). Educational assessment. In E.V. Hamayan and J.S. Damico, (eds.) *Limiting bias in the assessment of bilingual students.* (pp. 219-246). Austin, TX: Pro-Ed.

Collier, V.P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly, 21,*617-641.

Collier, V.P. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly, 23,* 509-531.

Council of Chief State School Officers. (1990). *School success for LEP students: The challenge and the state response.* Washington, D.C.: CCSSO, Resource Center on Educational Equity.

Council of Chief State School Officers. (1992a). *Recommendations* for *improving the assessment and monitoring of students with limited English proficiency.* Washington, D.C.: CCSSO, Resource Center on Educational Equity.

Council of Chief State School Officers. (1992b). *Summary of recommendations and policy implications for improving the assessment and monitoring of students with limited English proficiency* [On-line]. Available: http://www.ccsso.org

Cuevas, J. (1996, January). *Educating limited English proficient students.* (PD-96-01). San Francisco, CA: WestEd.

DelVecchio, A. & Guerrero, M. (1995, December). *Handbook of English language proficiency tests.* Albuquerque, NM: Evaluation Assistance Center–Western Region, New Mexico Highlands University.

Devine, J. (1988). The relationship between general language competence and second language reading proficiency: Implications for teaching. In R.C. Anderson & P.D. Pearson (Eds.). *Interactive approaches to second language reading* (pp. 260-277). New York: Cambridge University Press.

Duran, R.P. (1989). Assessment and instruction of at-risk Hispanic students. *Exceptional Children, 56* (2), 154-158.

Figueroa, R.A. (1990). Best practices in the assessment of bilingual children. In A. Thomas and J. Grimes (eds.). *Best practices in school psychology-II* (pp. 93-106). Washington, DC: National Association of School Psychologists.

Fradd, S.H., McGee, P.L., & Wilen, D.K. (1994). *Instructional assessment: An integrative approach to evaluating student performance.* Reading, MA: Addison Wesley.

Fuchs, D. & Fuchs, L. (1989). Effects of examiner familiarity on Black, Caucasian and Hispanic children: A meta-analyses. *Exceptional Children, 55* (4), 303-308.

Gandara, P. & Merino, B. (1993). Measuring the outcome of LEP programs: Test scores, exit rates and other mythological data. *Educational Evaluation and Policy Analysis, 15* (3), 320-338.

Garcia, G.E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly, 26* (4), 371-392.

Guerrero, M. & DelVecchio, A. (1996), March. *Handbook of Spanish language proficiency tests.* Albuquerque, NM: Evaluation Assistance Center–Western Region, New Mexico Highlands University.

Hamayan, E.V., & Damico, J.S. (Eds.) (1991). *Limiting Bias in the assessment of bilingual students.* Austin, TX: Pro-Ed.

Hodgkinson, H.L. (1992). *A demographic look at tomorrow.* Washington, D.C.: Institute for Educational Leadership, Center for Demographic Policy.

Hodgkinson, H.L. & Outtz, J.H. (1992). *The nation and the states: A profile and data book of America's diversity.* Washington, D.C.: Institute for Educational Leadership, Inc., Center for Demographic Policy.

Hoover, M.R., Politzer, R.L., & Taylor, O. (1987). Bias in reading tests for black language speakers: A sociolinguistic perspective. In A.G. Hilliard III (Ed.), *Testing African American students: Special reissue of the Negro Educational Review* (pp. 81-98). San Francisco, CA: Julian Richardson Association.

Huynh, H. (1990). Error rates in competency testing when test retaking is permitted *Journal of Educational Statistics, 15* (1), 39-52.

Ingels, S.J. (1993). *Strategies for including all students in national and state assessments: Lessons from a national longitudinal study.* Paper presented at the National Conference on Large Scale Assessment of the Council of Chief State School Officers, Albuquerque, NM.

Lacelle-Peterson, M. & Rivera, C. (1994, Spring). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review, 64* (1), 55-75.

Lam, T.C.M. (1993). Testability: A critical issue in testing language minority students with standardized achievement tests. *Measurement on in Counseling and Development, 26*, 179-191.

Lam, T.C.M. & Gordon, W.I. (1992, Winter). State policies for standardized achievement testing of limited English proficient students. *Educational Measurement: Issues and Practice, 11* (4), 18-20.

Liu, K., Thurlow, M., Vieburg, K., El Sawaf, H., & Ruhland, A. (1996). *Resources: Limited English proficient students in national and statewide assessments* (Minnesota Report 8). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

McDill, E.L., Natriello, G., & Pallas, A. (1985). Raising Standards and retaining students: The impact of the reform recommendations on potential dropouts. *Review of Educational Research, 55* (4), 415-433.

Mehrens, W. (1993). *Issues and recommendations regarding implementation of high school graduation tests*. Regional Policy Information Center. Oak Brook, IL: North Central Regional Educational Laboratory.

Miniccuci, C. & Olson, L.O. (1992, Spring). *Programs for secondary limited English proficient students: A California study* [On-line]. Available NCBE Focus: Occasional Papers in Bilingual Education, no. 5 : http://www.ncbe.gwu.edu/ncbepubs/focus/focus5.html

Minnesota Statutes. (1996). 126.262 [On-line]. Available: http://www.revisor.leg.state.mn.us/st96/126/262.html

National Academy of Education. (1996). *Quality and utility: The 1994 TSA in reading. The fourth report of the NAE panel on evaluation of the NAEP TSA: 1994 TSA in reading.* Stanford, CA: NAE, Stanford, University, School of Education, 53-73.

North Central Regional Educational Laboratory. (1996a). *Part I: Assessment of students with disabilities and LEP students. The status report of the assessment programs in the U.S. State student assessment programs database.* Oakbrook, IL: NCREL and CCSSO.

North Central Regional Educational Laboratory. (1996b). *The status of state student assessment programs in the United States: Annual report.* Oakbrook, IL: NCREL and CCSSO.

National Center for Education Statistics. (1992). *Language characteristics and academic achievement: A look at Asian and Hispanic eighth graders in NELS:88.* (NCES 92-479). Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement.

National Center for Education Statistics. (1996a). *Increasing the inclusion of students with disabilities and LEP students in NAEP* (Pre-publication copy). Washington, D.C.: NCES.

National Center for Education Statistics. (1996b). *Proceedings of the conference on inclusion guidelines and accommodations for LEP students in the LEP students in the NAEP: December 5-6, 1994* (NCES 96-86[1]). Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement.

National Research Council. (1997). *Improving schooling for language-minority children: A research agenda.* August, D. and Hakuta, K. (Eds.). Washington, D.C.: National Academy Press.

Navarette, C. & Gustkee, C. (1996). *A guide to performance assessments for linguistically diverse students.* Albuquerque, NM: Education Assistance Center–Western Region.

Nutall, E.V. (1987). Survey of current practices in the psychological assessment of limited English proficiency handicapped children. *Journal of School Psychology, 25*, 53-61.

O'Malley, J.M. & Valdez Pierce, L. (1994). State assessment policies, practices, and language minority students. *Educational Assessment, 2* (3), 213-255.

Oller Jr., J.W. & Damico, J.S. (1991). Theoretical considerations in the assessment of LEP students. In Hamayan, E.V. and Damico, J.S. (eds.). *Limiting bias in the assessment of bilingual students.* (pp. 77-110). Austin, TX: Pro-Ed.

Olmedo, E.L. (1981). Testing linguistic minorities. *American Psychologist, 36* (10), 1078-1085.

Olson, J. & Goldstein, A. (1996). *Increasing the inclusion of students with disabilities and LEP students in NAEP.* (Focus on NAEP Report–Pre-publication copy). Washington, D.C.: National Center for Education Statistics.

Penalosa, F. (1980). *Chicano sociolinguistics.* Rowley, MA: Newbury House Publishers.

Rivera, C. & Vincent, C. (1996). *High school graduation testing: Policies and practices in the assessment of LEP students.* Paper presented at the Council of Chief State School Officers, Phoenix.

Saville-Troike, M. (1991, Spring). *Teaching and testing for academic achievement: The role of language development* [On-line]. Available: NCBE Focus: Occasional Papers in Bilingual Education, no. 4: http://www.ncbe.gwu.edu/ncbepubs/focus/focus4.html

Shinn, M.R. & Tindal, G.A. (1988). Using student performance data in academics: A pragmatic and defensible approach to non-discriminatory assessment. In R.L. Jones (ed.), *Psychoeducational assessment of minority group children: A casebook* (pp. 383-407).

Spencer, B. (1991). *Eligibility/exclusion issues in the 1990 TSA.* Chicago, IL: Northwestern University, Methodology Research Center, NORC.

Spicuzza, R., Erickson, R., Thurlow, M., Liu, K., & Ruhland, A. (1996). *Input from the field on assessing students with limited English proficiency in Minnesota's Basic Requirements Exams* (Minnesota Report 2). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Taylor, O.L., & Lee, D.L. (1987). Standardized tests and African-American children: Communication and language issues. In A.G. Hilliard III (Ed.), *Testing African American students: Special reissue of the Negro Educational Review* (pp. 81-98). San Francisco, CA: Julian Richardson Association

Thurlow, M., Liu, K., Erickson, R., Spicuzza, R., & El Sawaf, H. (1996). *Accommodations for students with limited English proficiency: Analysis of guidelines from states with graduation exams* (Minnesota Report 6). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

United States General Accounting Office. (1994, January). *Limited English proficiency: A growing and costly educational challenge facing many school districts.* Report to the chairman, Committee on Labor and Human Resources, U.S. Senate, Washington, D.C., USGAO. GAO/HEHS 94-38.

Walstad, W.B. (1984). Analyzing minimal competency test performance. *Journal of Educational Research,* 77(5), 261-266.

Wilde, J. & Sockey, S. (1995). *Evaluation handbook.* Albuquerque, NM: Evaluation Assistance Center–Western Region.

Wilen, D.K. & van Maanen Sweeting, C. (1986). Assessment of LEP Hispanic Students. *School Psychology Review, 15* (1), 59-75.

Zehler, A.M., Hopstock, P.J., Fleischman, H.L., & Greniuk, C. (1994, March 28). *An examination of assessment of limited English proficiency students* [On-line].*Task Order D070 Report.* Arlington, VA: Special Analysis Center. Available: http://www.ncbe.gwu.edu/miscpubs/siac/lepasses.html

Zlatos, B. (1994, November). Don't test, don't tell. *American School Board Journal,* 24-28.

Zlatos, B. (1995, January). Scores that don't add up. *Network News & Views,* 100-101.

NATIONAL CENTER ON EDUCATIONAL OUTCOMES

University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612/626-1530 • Fax 612/624-0879
http://www.coled.umn.edu/NCEO

The College of Education & Human Development
UNIVERSITY OF MINNESOTA